

# Elementary test for non-classicality based on measurements of position and momentum

Luca Frezza,<sup>1</sup> Johannes Borregaard,<sup>2,3</sup> and Anders S. Sørensen<sup>2</sup>

<sup>1</sup>*Dipartimento di Fisica, Università degli Studi di Milano, I-20133 Milano, Italy*

<sup>2</sup>*The Niels Bohr Institute, University of Copenhagen,  
Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark*

<sup>3</sup>*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*

(Dated: 04 August 2015)

We generalise a non-classicality test described by Kot *et al.* [Phys. Rev. Lett. **108**, 233601 (2010)], which can be used to rule out any classical description of a physical system. The test is based on measurements of quadrature operators and works by proving a contradiction with the classical description in terms of a probability distribution in phase space. As opposed to the previous work, we generalise the test to include states without rotational symmetry in phase space. Furthermore, we compare the performance of the non-classicality test with classical tomography methods based on the inverse Radon transform, which can also be used to establish the quantum nature of a physical system. In particular, we consider a non-classicality test based on the so-called filtered back-projection formula. We show that the general non-classicality test is conceptually simpler, requires less assumptions on the system and is statistically more reliable than the tests based on the filtered back-projection formula. As a specific example, we derive the optimal test for a quadrature squeezed single photon state and show that the efficiency of the test does not change with the degree of squeezing.

PACS numbers: 42.50.Dv

## I. INTRODUCTION

Nearly a century after its foundation [1, 2], Quantum Mechanics is a well established theory for microscopic and mesoscopic phenomena. Nonetheless, the logical structure of quantum mechanics is still challenging our perception of the theory, spurring investigations of a deeper understanding of the quantum theory and its relation to the classical perception of reality [3, 4]. In particular, the conceptual differences between the quantum and classical world remain a fundamental topic of investigation, in part due to the potential applications of quantum phenomena for quantum information technology [5]. To this end, experiments have sought to demonstrate effects which are truly *non-classical*, but the challenge in this case is how to prove that something really is non-classical. Clearly, the mere consistency with predictions from quantum theory is not sufficient to claim the observation of genuine quantum effects as this does not rule out a classical description of the same phenomena, although the classical explanation may seem different from the quantum mechanical. On the contrary, throughout this article, we shall use the definition that an experimental observation of a *genuine non-classical* effect should only assume concepts from classical physics and then demonstrate that these assumptions lead to a contradiction with the observed experimental result. In other words, the goal of the non-classicality test is to convince a physicist trained in classical physics, but completely ignorant of quantum theory, that his/her perception of the world is incorrect. A non-classicality test fulfilling this requirement based on measurements of quadrature operators, *e.g.*, position  $x$  and momentum  $p$ , was presented in Ref. [6]. Further-

more, the test was used to demonstrate that the electric and magnetic field of a single photon state, cannot be described classically. In this article, we generalise the approach of Ref. [6] and develop the theory for genuine non-classicality tests based on measurements of quadrature operators for a physical system. As opposed to Ref. [6], which only considered rotationally symmetric states, we consider states of arbitrary shape. Furthermore, we compare the obtained results with more conventional approaches relying on tomographic reconstructions, *e.g.* inverse Radon transformations, and demonstrate that the tests developed here are both conceptually simpler, require fewer assumptions on the system, and are statistically more reliable.

One of the most impressive examples of a genuine non-classical effect is the violation of Bell's inequality. In Bell's inequality, two simple classical assumptions are made, locality and realism, and this is shown to lead to a contradiction with the observed experimental results. Remarkably, this can be done without any assumptions about the physics of the underlying physical system. On the other hand, Bell's inequality is often quite difficult to violate in practice [7, 8]. We, therefore, take a different approach and consider non-classicality tests based on measurements of quadrature quantities of the form

$$Q_\theta \equiv \cos \theta x + \sin \theta p \quad (1)$$

for a single system. We shall assume that  $Q_\theta$  exist for any  $\theta \in [0, \pi]$ , in the sense that they are well defined measurable observables of the physical system. In particular, we are assuming that there exist two independent observables  $x$  and  $p$ , which might be canonical variables although this is not necessary. As considered in Ref. [6],

this is for example the situation for optical fields measured by homodyne detection. In this case, a completely classical analysis reveals that  $Q_\theta$  are the measured observables with the angle  $\theta$  controlled by the phase of an interfering laser field and  $x$  and  $p$  are, *e.g.*, the electric and magnetic field components of a single mode of the measured optical field. Another example can be found in the emerging field of opto-mechanics [9, 10] where a lot of effort is devoted to proving that mechanical oscillators can have non-classical behaviour. Here, a measurement of the position  $x_m(t)$  at a time  $t$  will have the form in Eq. (1) with  $\theta = \omega_m t$  where  $\omega_m$  is the vibration frequency of the oscillator and  $x$  and  $p$  are the (suitably rescaled) position and momentum operators at time  $t = 0$ . Finally, operators of the form in Eq. (1) can also be measured for atomic ensembles where  $x$  and  $p$  represents two different spin components of the ensemble [11]. In principle, operators of the form in Eq. (1) can be used to show a violation of Bell's inequality [12] but this requires preparing quantum states, which are highly difficult to produce. To avoid this, we make a stronger assumption compared to Bell's inequalities and assume that the measured observables are described by the relation in Eq. (1) as prescribed by the underlying physics of the system. This assumption makes our non-classicality inherently weaker than Bell's inequality, but on the other hand enables the violation of the non-classicality test under less stringent experimental requirements.

As a second assumption underlying *any* classical description of the system, we assume that the quadratures defined by Eq. (1) commute with one another. As a result, these two hypotheses give the complete picture of the observable quantities that we will need for our non-classicality test. Mathematically speaking, it means that we postulate the algebra of observables to be the commutative algebra generated by the rotated quadratures. These hypotheses underlie the classical description of the system, which we will be able to rule out by deriving a contradiction from experimental data. In this description where  $x$  and  $p$  are commuting quantities, the state of the system will be characterized by a joint  $xp$ -probability distribution, also called the phase space distribution  $W_c(x, p)$ . Then, the contradiction consists in showing that such a distribution does not exist for the measured experimental outcomes.

In quantum mechanics, it is well known that there does not exist a joint probability distribution for  $x$  and  $p$  since the position and momentum of a particle do not commute. The closest thing that one can have is the Wigner function  $W(x, p)$  [13], which correctly predicts the probability outcome of measurements of the observables  $Q_\theta$  but which cannot be considered a proper distribution function since it can attain negative values (see Fig. 1). The negativity of the Wigner function will be the underlying quantum mechanical reason for the failure of the classical description in the non-classicality test derived below. In quantum mechanics, the Wigner function is, however, not the only possible phase space distribution and nu-

merous others have been considered as the basis for non-classicality criteria. For instance the  $P$ -function represents the expansion of a field on coherent states and is thus negative for any state with an uncertainty squeezed below the level of a coherent state. Since, according to quantum theory, classical light sources produce coherent states, any state produced by classical sources will have a positive  $P$ -function. As a consequence, states with negative  $P$ -functions are often referred to as non-classical and based on this definition numerous observable non-classicality criteria have been derived [14–17]. These criteria, however, inherently rely on quantum mechanics since the underlying criteria is based on having smaller uncertainty than expected from the quantum mechanical uncertainty relation. With the definition used here, *i.e.*, that a non-classicality criteria should only rely on concepts from classical physics, tests based on the negativity of the  $P$ -function are thus not applicable. From a mathematical perspective, however, the test that we will derive is highly related. In particular our test is essentially the same as the one derived for the  $P$ -function in Refs. [14, 17], but it is translated to Wigner functions in such a way that quantum mechanics is not assumed, thus leading to a stronger test.

## II. SIMPLE NON-CLASSICALITY TEST

### A. Premise

Before proceeding, we shortly review the non-classicality test presented in Ref. [6], which is based on previous work by Bednoraz and Belzig [18]. We will emphasise the fundamental ideas behind the test and outline how it can be generalized into the more general test introduced below.

As already discussed in the introduction, a genuine non-classicality criterion should directly disprove the classical description without logically relying on quantum mechanics. Because the state of the system is classically characterized by a probability distribution  $W_c(x, p)$  over the phase space of  $x$  and  $p$ , the expectation value of any function  $\mathcal{A}(x, p)$  is

$$\langle \mathcal{A} \rangle \equiv \int \int dx dp W_c(x, p) \mathcal{A}(x, p). \quad (2)$$

Mathematically speaking, Eq. (2) is the integral representation of positive, normalized, linear functionals on the  $C^*$ -algebra generated by the observables. Accordingly, any non-negative quantity  $\mathcal{F}(x, p) \geq 0$  should have a non-negative expectation value,

$$\langle \mathcal{F} \rangle \equiv \int \int dx dp W_c(x, p) \mathcal{F}(x, p) \geq 0, \quad (3)$$

because the integrand is non-negative,  $W_c(x, p)$  being a probability distribution. The violation of Eq. (3) disproves  $W_c(x, p)$  as a probability distribution and thus

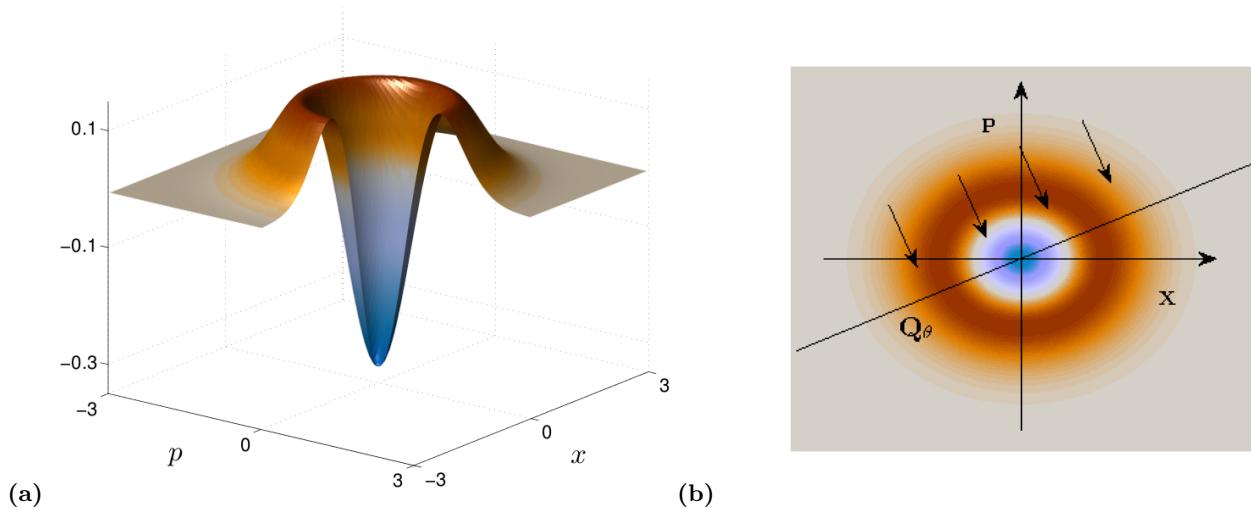


FIG. 1 (colour online). (a) Wigner function  $W(x, p)$  of the single photon state of a single mode of light. It is not a probability distribution as it should be in the classical framework, since it takes on negative values. (b) Top view of the same function. The colour gradient allows to identify the negative region in the centre. The cut represented by the line is the quadrature  $Q_\theta$ . The marginal distribution corresponding to the probability distribution for measurement outcomes for  $Q_\theta$  is obtained by integration of  $W(x, p)$  along the perpendicular directions to the cut, as shown by the arrows.

rules out any classical description of the system. We cannot, however, directly determine  $\langle \mathcal{F} \rangle$  since we cannot measure  $x$  and  $p$  simultaneously and we have to infer it from measurements of  $Q_\theta$ .

In Ref. [6], the test function  $\mathcal{F}$  was chosen to be the square of a polynomial in the squared radius  $r^2 = x^2 + p^2$  since the underlying physical state was assumed to be rotationally symmetric. When the polynomial in  $r$  was expanded, the various moment of  $r$  could be expressed in moments of the rotated quadratures  $Q_\theta$ . Consequently, the mean value  $\langle \mathcal{F} \rangle$  together with its experimental error were directly computable from the experimental outcomes. In Ref. [6], the developed test was applied to the experimental data for an approximate single photon state, attaining the violation of Eq. (3) by nineteen standard deviations [6].

## B. General formulation

We will now generalise the test function considered in Ref. [6] so that it does not rely on rotational symmetry. Consequently, the generalized formulation can exploit the freedom in the choice of phases for the quadratures to construct an optimal test function for an arbitrary quantum state.

Let us define the problem in more rigorous terms. To formulate a test, we need a suitable class of non-negative test functions  $\mathcal{F}$  defined on the phase space and a method to determine  $\langle \mathcal{F} \rangle$ , together with its empirical error, directly from the measured quadrature moments. The general class of smooth non-negative functions on the phase space is too large for our purposes and we therefore restrict the test to non-negative polynomials in the vari-

ables  $x$  and  $p$ . Note, that such polynomials can be peaked in a region around the negativity of  $W(x, p)$  and since  $W(x, p)$  typically vanishes faster than any power law outside its significant support, the divergent terms of the polynomials are irrelevant (see Fig. 2). Consequently, we do not lose much generality by restricting to polynomials, and as we will show below, it allows us to make an explicit representation of the test function in terms of quadrature quantities.

Motivated by these considerations, we consider test functions of the form

$$\mathcal{F}(x, p) = \left( 1 + \sum_{i=1}^N \sum_{k=0}^i D_{i;k} x^{i-k} p^k \right)^2. \quad (4)$$

Here we have, for simplicity, put the zeroth order coefficient equal to one, since it is irrelevant to the effectiveness of the test as discussed below. Furthermore, we choose to ensure the positivity by only considering squared polynomials. This choice simplifies the test as discussed in Appendix A, but does not exhaust all the possible, non-negative polynomials. This is related to the 17<sup>th</sup> Hilbert problem [17, 19]. However, since a polynomial of sufficiently high order can approximate an analytical function sharply peaked at any particular value where  $W(x, p)$  is negative, we believe that considering this restricted set does not degrade the efficiency of the test significantly.

To have an experimentally applicable test,  $\mathcal{F}(x, p)$  needs to be expressed in terms of powers of measurable quantities. For a fixed  $i$ ,  $Q_\theta^i$  directly gives a sum of  $(i+1)$  moments of the form  $x^{i-k} p^k$  with  $k = 0, \dots, i$ . Hence, if one measures  $m \geq i+1$  different angles  $\theta_j$ , one can always combine those linear equations by simple linear algebra

to arrive at

$$x^{i-k}p^k = \sum_{j=1}^m T_{i;k,j} Q_{\theta_j}^i, \quad (5)$$

where  $T_{i;k,j}$  are real-valued coefficients which have to satisfy the linear constraint

$$\sum_{j=1}^m \binom{i}{k} \cos^i \theta_j \tan^k \theta_j T_{i;k',j} = \delta_{k,k'} \quad \forall i. \quad (6)$$

Note that at least  $(i+1)$  linearly independent quadratures are needed to express all  $(i+1)$  moments of  $x$  and  $p$ , but if  $m > i+1$ , the constraint does not give a unique solution and there is a freedom in the choice of the  $T_{i;k,j}$  coefficients.

Before proceeding, we stress that we have assumed commutativity of  $x$  and  $p$  in attaining Eq. (5), *i.e.*, we assume a classical description of the quadratures in terms of  $x$  and  $p$ . From Eq. (5), we can express the test function in terms of the quadratures  $Q_{\theta}^i$ , which are directly measurable by assumption. Accordingly, we have explicitly shown that the test function  $\mathcal{F}(x, p)$  is in the algebra of observables and through Eq. (5), we are able to compute its mean value. Finally, it is worth noting that this strikingly simple linear relation is the very heart of the procedure we consider. This is the trick that allows us to avoid the complicated machinery of the inverse Radon algorithms for the inverse problem of finding  $W(x, p)$  from the measured quadratures  $Q_{\theta}$  as we will discuss in Sec. III.

In summary, once the polynomial order of the test function is fixed to  $2N$ , the test is characterized by three sets of parameters:

- the set of free coefficients  $D_{i;k}$  that define the test function over the phase space,
- the set of  $m \geq 2N+1$  free phases  $\{\theta_j\}$  which identify the quadratures we measure,
- the set of coefficients  $T_{i;k,j}$ , which allow to express the  $x^{i-k}p^k$  moments in terms of  $Q_{\theta_j}$ . These are not completely free to vary but has to satisfy the linear constraint in Eq. (6).

We will formally refer to a test,  $\mathbb{T}$ , as the string of parameters

$$\mathbb{T} \equiv (N, D_{i;k}, \theta_j, T_{i;k,j}). \quad (7)$$

We will now describe, in more details, how to express the expectation value of the test function and its relative empirical error in terms of the measurable quadratures  $Q_{\theta_j}$ . Expanding the square in the test function in Eq. (4)

and using Eq. (5), we get

$$\begin{aligned} \mathcal{F}(x, p) &= 1 + \sum_{i=1}^{2N} \sum_{k=0}^i C_{i;k} x^{i-k} p^k \\ &= 1 + \sum_{j=1}^m \sum_{i=1}^{2N} Q_{\theta_j}^i \left( \sum_{k=0}^i C_{i;k} T_{i;k,j} \right) \\ &= 1 + \sum_{j=1}^m \mathcal{H}_j(Q_{\theta_j}). \end{aligned} \quad (8)$$

For simplicity, we have here introduced the coefficients  $C_{i;k} \equiv 2D_{i;k} + \sum_{a+a'=i} \sum_{b+b'=k} D_{a;b} D_{a';b'}$ , where  $D_{i;k}$  vanishes for  $i > N$ . Furthermore, we have introduced the set of  $m$  polynomials,  $\mathcal{H}_j(Q_{\theta_j})$ , (one for each phase cut) defined as

$$\mathcal{H}_j(Q_{\theta_j}) \equiv \sum_{i=1}^{2N} Q_{\theta_j}^i \left( \sum_{k=0}^i C_{i;k} T_{i;k,j} \right). \quad (9)$$

Eq. (8) shows that according to the relation in Eq. (5), the test function can be expressed as a polynomial of quadratures, *i.e.*, as a sum of  $2N+1$  polynomials of degree  $2N$ , one for each independent phase cut.

To evaluate the expectation value of the final expression in Eq. (8), we consider each quadrature as an independent random variable distributed according to its marginal. In principle, if one has a very large number of cuts (in particular if the number of cuts  $m$  exceeds the minimal required value  $m \geq 2N+1$ ), one could argue that nearby cuts will have similar moments and this may be used in a statistical analysis to reduce the variance on the experimental data. On the other hand, we prefer to make as few assumptions as possible and therefore consider each cut to be an independent variable, distributed according to its marginal. Thus, for fixed phase cut, the expectation value  $\langle \mathcal{H}_j(Q_{\theta}) \rangle$  is formally given by integration with the marginal distribution and is experimentally estimated as

$$\langle \mathcal{H}_j(Q_{\theta}) \rangle \simeq \frac{\sum_{n=1}^{M_j} \mathcal{H}([Q_{\theta}]_n)}{M_j}, \quad (10)$$

with  $[Q_{\theta}]_n$  being the  $n$ -th outcome within the dataset consisting of a total of  $M_j$  measurements of the random variable  $Q_{\theta}$ . Accordingly, we can compute the average value of the test function  $\mathcal{F}$  as

$$\langle \mathcal{F} \rangle = 1 + \sum_{j=1}^m \langle \mathcal{H}_j(Q_{\theta_j}) \rangle, \quad (11)$$

which according to a classical description, must be positive (See Eq. (3)).

Quantum mechanically,  $\langle \mathcal{F} \rangle$  is given by Eq. (2) with  $W_c(x, p)$  being the Wigner function  $W(x, p)$  describing the phase-space distribution. For a general density matrix,  $\hat{\rho}$ , the Wigner function is defined as

$$W(x, p) = \frac{1}{\hbar\pi} \int_{-\infty}^{\infty} \langle x + y | \hat{\rho} | x - y \rangle e^{-2ipy/\hbar} dy, \quad (12)$$

where  $|x \pm y\rangle$  are eigenstates of the position operator with eigenvalue  $x \pm y$ . The Wigner function is often referred to as a quasi-probability distribution since it can attain negative values and is thus not a true probability distribution. This possibility of the Wigner function being negative is the key aspect of our non-classicality test. With a negative Wigner function, Eq. (2) and the equivalent form in Eq. (11) are no longer guaranteed to be positive leading to contradiction with classical physics. Given the Wigner function describing a quantum system, we can find the set of coefficients  $D_{i;k}$  that minimizes  $\langle \mathcal{F} \rangle$ , by solving a linear optimization problem as described in Ref. [6]. In an experimental test, however, we also have to consider the empirical error of our estimate of  $\langle \mathcal{F} \rangle$ . Since each random variable,  $Q_\theta$ , is treated as statistically independent for different  $\theta$ , we can estimate the variance of  $\langle \mathcal{F} \rangle$  as

$$\sigma_{\mathcal{F}}^2 = \sum_{j=1}^m \frac{1}{M_c - 1} \left( \langle \mathcal{H}_j^2(Q_\theta) \rangle - \langle \mathcal{H}_j(Q_\theta) \rangle^2 \right), \quad (13)$$

where we have assumed for simplicity that the same number of measurements  $M_c$  is performed for each cut. This is sensible because of the statistical independence assumed, which means that no bias is justified (for a formal justification of this choice see Appendix A).

We are interested in classifying how well the various non-classicality tests perform. To do so, we introduce the following quantity for any non-classicality test,

$$\mathcal{G}[\mathbb{T}] \equiv \lim_{M \rightarrow \infty} \frac{\langle \mathcal{F} \rangle}{\sigma_{\mathcal{F}} \sqrt{M - m}}, \quad (14)$$

which depends on all the characteristics of the particular test  $\mathbb{T}$ , but not on the total number of measurements  $M \equiv m M_c$  because  $\sigma_{\mathcal{F}}$  scales asymptotically as  $1/\sqrt{M}$ . In fact, as shown in Appendix A,  $\sigma_{\mathcal{F}}$  goes exactly as  $1/\sqrt{M - m}$  for the optimal test and the  $M$  dependence thus disappears even before taking the limit. Note that because  $\mathcal{G}$  involves a ratio, any overall proportionality constant in the test function is irrelevant. We have therefore chosen to set the first coefficient in the polynomial in Eq. (4) equal to unity to get rid of any overall scaling factor.

The optimal test among all possible tests  $\mathbb{T}$  (See Eq. (7)) is defined as the one, which minimizes  $\mathcal{G}$  in Eq. (14). Formally, the search for the optimal test involves an algorithm to choose a suitable test among a class of logically equivalent tests. Such an optimization is generally hard to do analytically and we will therefore treat it numerically. Since we want to prove the non-classicality of the system, it is natural to optimize the test using the quantum mechanical description of its state, *i.e.*, by assuming the suitable Wigner function to be the correct description of the unknown distribution  $W(x, p)$ . We can then employ simple symmetry considerations about  $W(x, p)$  in the optimization, as we show below for rotational invariant states and squeezed states.

Note, however, that even though the optimization assumes the quantum mechanical description of the state, the test itself is still completely classical, and the assumption of quantum mechanics at this stage does not decrease the strength of the test. In an experiment, the optimal choice of phase cuts has to be determined before the experimental measurements while the optimization of the other free parameters can be performed afterwards and may be optimized over the experimental data. If this is done, the statistical analysis of the experiments should also include the freedom in varying these parameters. For simplicity, we will just assume that one uses the coefficients expected from theory, in which case we do not have to worry about this complication.

### C. Rotationally invariant states

If the distribution  $W(x, p)$ , describing the state, is rotationally invariant, we naively expect that all the parameters of the optimal test function should exhibit this rotationally invariance, *e.g.*, that the optimal phase cuts  $\{\theta_j\}$  should be uniformly distributed. This was assumed in Ref. [6] but not formally shown. We will now show that this is indeed the case.

We explicitly define the rotational version of the test function as

$$\mathcal{F}(x, p) = \left( 1 + \sum_{i=1}^{\lfloor N/2 \rfloor} d_i r^{2i} \right)^2, \quad (15)$$

with  $r^2 = x^2 + p^2$  being the squared radius in the phase plane. Apart from the fact that we have formally reduced the set of coefficients  $D_{i;k}$  to the lighter set of  $d_i$ , we make a further reduction when we relate the radial moments to the quadrature moments in a similar way as done in Eq. (5). We seek a linear relation of the form

$$r^{2i} = \sum_{j=0}^m t_{i;j} Q_{\theta_j}^{2i}, \quad (16)$$

with the set of coefficients,  $t_{i;j}$ , being the analogue of the set of  $T_{i;k,j}$  in Eq. (5). The existence of such a linear relation, provided a sufficient number of phase cuts, can be deduced from arguments similar to the ones leading to the relation in Eq. (5). Note that odd powers of the quadratures do not appear in the above formula, since the parity symmetry of  $r^{2i}$  guarantees the null contribution of such terms. Denoting rotationally invariant tests as  $\mathbb{T}_{rad}$ , we prove in Appendix A the following theorem.

**Theorem 1.** *Whenever  $W(x, p)$  is rotationally symmetric, and takes on negative values, the optimal test, *i.e.*, the test  $\mathbb{T}$  that minimizes  $\mathcal{G}$ , is a radial test  $\mathbb{T}_{rad}$  and the*

cuts are chosen in such a way that

$$\theta_j = \frac{\pi j}{m} \quad j = 0, \dots, m-1 \quad (17)$$

$$t_{i;j} \equiv t_i = \left( \frac{2i}{i} \right)^{-1} \frac{4^i}{m} \quad \forall j. \quad (18)$$

Furthermore  $\mathcal{G}$  is independent of the number of cuts provided that  $m \geq N + 1$ .

The theorem states that the optimal test function is a radial polynomial and the cuts should be distributed uniformly over the phase space whenever  $W(x, p)$  is rotationally symmetric. Note that the numerical optimization of the test is a lot simpler than in the general case because the optimal test function can be put into the form in Eq. (15), which has fewer free parameters. Yet, the reader should bear in mind that with the choice given by Eqs. (17) and (18), the linear relation in Eq. (16) holds and can be applied even for non-symmetric states. Hence, the assumptions made can never lead to a state incorrectly being assigned as non-classical even if the assumptions are not fulfilled. Applying a rotationally invariant test to a non rotationally invariant state will thus only lead to a less efficient test but not to an incorrect result.

For the test  $T_{rad}$ , we can express the test function, using Eqs. (16) and (18), as

$$\mathcal{F} = 1 + \sum_{j=0}^m \mathcal{H}(Q_{\theta_j}), \quad (19)$$

where  $\mathcal{H}$  is now the same polynomial for all the phases, contrary to the general case in Eq. (11). Defining  $c_i = 2d_i + \sum_{a+a'=i} d_a d_{a'}$ , we can explicitly express this polynomial as

$$\mathcal{H}(Q_{\theta}) \equiv \sum_{i=1}^{2N} c_i t_i Q_{\theta}^{2i}. \quad (20)$$

Otherwise all considerations about the evaluation of  $\mathcal{G}$  are the same as for the general test. Most importantly, we treat every cut as statistically independent, even in the case of actual rotational symmetry, because we do not want to assume this symmetry.

For the single photon state considered in Ref. [6], which corresponds to the first excited state of a harmonic oscillator, the Wigner function is

$$W(x, p) = \frac{1}{\pi} \left( 2(x^2 + p^2) - 1 \right) \exp[-(x^2 + p^2)]. \quad (21)$$

We have numerically optimized  $\mathcal{G}$  with respect to  $\mathcal{T}$  for increasing polynomial order. Fig. 2 shows the optimal  $\mathcal{G}$  as a function of  $N$ , the order of the test function being  $2N$ . We find a violation of Eq. (3) for  $N \geq 4$  in agreement with what was found in Ref. [6]. Furthermore, we see that the graph plotted in Fig. 2 exhibits an asymptotic saturation of  $\mathcal{G}$ . Accordingly, in the case of arbitrarily high number

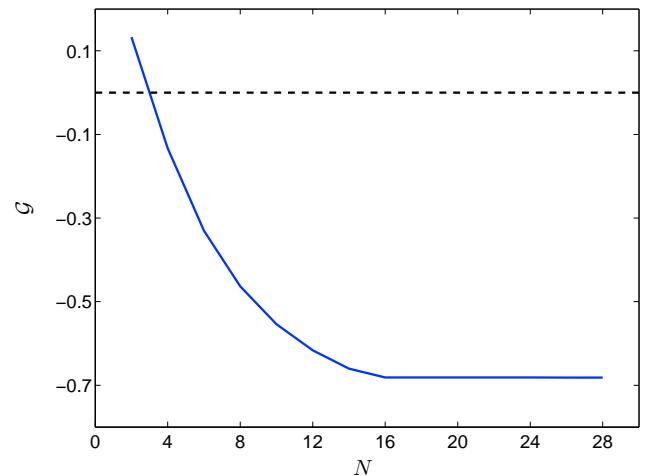


FIG. 2 (colour online). Violation of non-classicality for the first excited state of a harmonic oscillator, e.g. a single photon state. The figure shows the optimal expectation value of the test function relative to its standard deviation ( $\mathcal{G}$ ) as a function of the order of the polynomial in Eq. (15). Experimentally this may result in a violation of classicality by  $-\sqrt{M - m}\mathcal{G}$  standard deviations, with  $M$  being the total number of measurements and  $m \geq N + 1$  the number of measured quadratures.

of measurements, increasing the polynomial degree beyond  $N \sim 16$  gives more degrees of freedom for the test but does not decrease  $\mathcal{G}$  significantly. This demonstrates that polynomial test functions are efficient. In principle, more general, bounded and smooth test functions could be employed but the results obtained here suggest that this is not necessary.

While  $\mathcal{G}$  quantifies the relation between different tests, it does not directly give the statistical strength,  $\langle \mathcal{F} \rangle / \sigma_{\mathcal{F}}$  of a test. The latter can, however, be obtained by simply multiplying  $\mathcal{G}$  with  $\sqrt{M - m}$ . As stated above  $\mathcal{G}$  is independent of the number of cuts, but because of the dependence of  $m$  in  $\sqrt{M - m}$  it will in general be better to have the least number of cuts, *i.e.*,  $m = N + 1$ , although this dependency will be small if a large number of measurement are performed  $M \gg m$ . In order to keep  $m$  as small as possible we thus find from Fig. 2 that  $\langle \mathcal{F} \rangle / \sigma_{\mathcal{F}}$  is minimised by working around  $N = 16$  although higher order tests  $N > 16$  will produce essentially similar results if sufficiently many measurements are performed  $M \gg m$ . For imperfectly prepared states, however, the area in phase space where the Wigner function is negative is smaller. Hence in this case it may be desirable to work at a higher  $N$  where the function can be tailored to be more narrow [6]. In Fig. 3, we plot the optimal test function for the minimum argument  $N = 16$ . It is seen that the test function is peaked at the negative region of the Wigner function and that the tails of the polynomial are outside the significant support of  $W(x, p)$  and are thus not affecting the test.

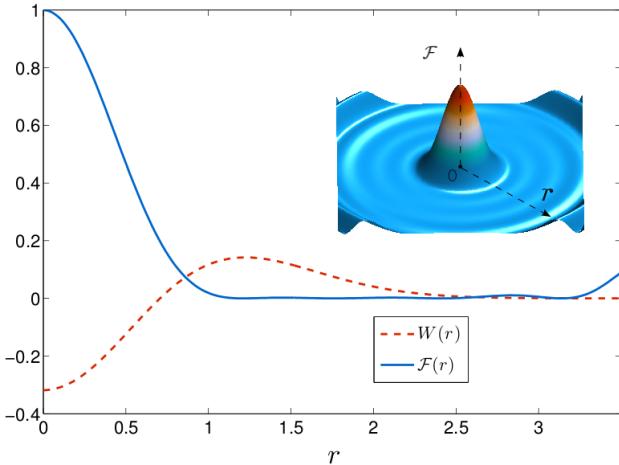


FIG. 3 (colour online). Radial profile of  $W(x, p)$  for the single photon state (red dashed curve) and of the optimal test function  $\mathcal{F}(x, p)$  (light blue solid curve) for  $N = 16$ . The inset shows the corresponding 3D plot of  $\mathcal{F}(x, p)$ .

#### D. Squeezed states

We now turn to the case where  $W(x, p)$  is not rotationally symmetric. As a specific example, we consider a squeezed single photon state and show that the optimal test function differs significantly from a rotationally invariant test. Let  $W(x, p)$  be the Wigner function of the single photon state. The Wigner function,  $W_\lambda(x, p)$  of the squeezed state is given by the rescaling

$$W_\lambda(x, p) = W(\lambda x, p/\lambda), \quad (22)$$

where  $\lambda$  characterizes the amount of squeezing in either  $x$  ( $\lambda > 1$ ) or  $p$  ( $0 < \lambda < 1$ ). The unsqueezed state corresponds to  $\lambda = 1$ . Fig. 4 shows the result of the squeezing transformation; the Wigner function of the single photon state is stretched into an ellipse for the squeezed state.

The squeezed state still exhibit elliptical symmetry but the marginal distributions are no longer identical as they were for the rotationally invariant state. The extra freedom in the choice of cuts present in the general test is therefore of importance. As an ansatz for a test function, we consider

$$\mathcal{F}_\lambda(x, p) = \left( 1 + \sum_{i=1}^{\lfloor N/2 \rfloor} d_i r_\lambda^{2i} \right)^2, \quad (23)$$

with  $r_\lambda^2 = (\lambda x)^2 + (p/\lambda)^2$  being the rescaled radius in the phase plane. Eq. (23) is a rescaling of the radial functions in Eq. (15) by the squeezing parameter,  $\lambda$ . Note that with this type of test function, the average value of  $\mathcal{F}_\lambda$  with a phase space distribution  $W_\lambda$  is independent of  $\lambda$ . The polynomial expression of  $\mathcal{F}_\lambda(x, p)$  with respect to the rotated quadratures is obtained in a straightforward manner from Eq. (16) with the substitution  $x \rightarrow \lambda x$  and

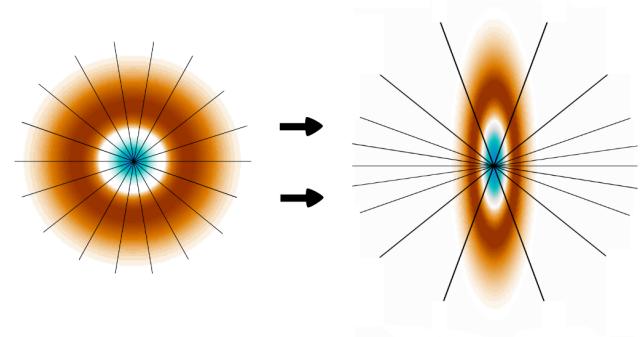


FIG. 4 (colour online). Sketch (not to scale) of  $W(x, p)$  (left) and  $W_\lambda(x, p)$  (right), together with a representation of the optimal distribution of phase cuts. The cuts are uniformly distributed for the single photon state and stretched along the  $x$  axis for the squeezed state. Note that the cuts stretch along the opposite direction of stretching for  $W_\lambda(x, p)$ . This reflects that most cuts are placed at angles where the distribution varies the fastest. The inner/blue region is where the distribution takes on negative values.

$p \rightarrow p/\lambda$  such that

$$\begin{aligned} r_\lambda^{2i} &= \sum_{j=0}^m t_{i;j} (\cos \theta_j \lambda x + \sin \theta_j p/\lambda)^{2i} \\ &= \sum_{j=0}^m t_{i;j} (\lambda \cos \theta_j)^{2i} \left( x + \frac{\tan \theta_j}{\lambda^2} p \right)^{2i} \\ &= \sum_{j=0}^m t_{i;j} \left( \lambda \frac{\cos \theta_j}{\cos \tilde{\theta}_j} \right)^{2i} Q_{\tilde{\theta}_j}^{2i} \end{aligned} \quad (24)$$

where  $\tilde{\theta} \in [0, \pi)$  satisfies  $\tan \tilde{\theta}_j = \frac{\tan \theta_j}{\lambda^2}$ . Geometrically, this transformation  $\theta \rightarrow \tilde{\theta}$  is the polar representation of the rescaling that we have performed on the phase plane.

In Appendix A, we show that not only is the average value of  $\mathcal{F}_\lambda$  independent of  $\lambda$ , but also the same is true for the figure of merit  $\mathcal{G}$ . This is formally stated in the following theorem.

**Theorem 2.** (Invariance under squeezing) *The minimum of  $\mathcal{G}$  for the squeezed state is independent of the squeezing parameter  $\lambda$  and equal to the minimum for the rotationally symmetric state. The optimal test function is given by Eq. (23) with  $d_i$  being the same as the optimal ones for the single photon state. Furthermore  $\theta_j$  and  $t_{i;j}$  given in Eqs. (17,18) transform to the “rescaled” values*

$$\begin{aligned} \theta_j &\rightarrow \tilde{\theta}_j \\ t_{i;j} &\rightarrow \tilde{t}_{i;j} \equiv t_{i;j} \left( \lambda \frac{\cos \theta_j}{\cos \tilde{\theta}_j} \right)^{2i} \end{aligned}$$

In Appendix A, the theorem is stated in a more general fashion along with the proof of it. In other words, allowing the distribution of phase cuts to deviate from the uniform distribution results in non-classicality tests, which are equally efficient for the squeezed and unsqueezed single photon states. The introduction of the squeezing thus does not degrade the test and the results of Fig. 2 are also applicable to the squeezed single photon state. This is in contrast with the reconstruction method based on standard inverse Radon transformation, where squeezing would introduce additional ambiguities and/or a reduction in the statistical significance of the results as discussed below.

### III. INVERSE RADON TRANSFORM

#### A. Introduction

Above, we have presented a simple and general non-classicality test that detects negativity of a phase space distribution  $W(x, p)$  resulting in the breakdown of any classical description of the system. This objective could, however, also be achieved through standard state reconstruction methods based on the inverse Radon transform. As already stated in the previous section, we cannot directly measure the distribution  $W(x, p)$  with arbitrary precision. What we can measure is the probability distribution of the quadrature operators  $Q_\theta$ , which is given by the marginal distribution known as the Radon transform,

$$\mathbf{R}W(\theta, s) \equiv \int \int_{\mathbb{R}^2} W(x, p) \delta(s - x \cos \theta - p \sin \theta) dx dp. \quad (25)$$

Formally,  $\mathbf{R}W(\theta, s)$  is an integral transform obtained from integration along the rays perpendicular to the cut characterized by  $\theta$  at distance  $s$  from the origin. Fig. 1(b) shows graphically how this transform works. The inverse process is referred to as the inverse Radon transformation, which allows to extract the properties of the underlying distribution  $W(x, p)$  from  $\mathbf{R}W(\theta, s)$ . We will now consider the use of the inverse Radon transform to show non-classicality. Furthermore, we will compare it to the simple test discussed above in order to determine advantages and disadvantages of the two methods for an experimental demonstration of non-classicality.

The inversion of the Radon transform is not a simple task, but because of its importance in image reconstruction, there is a wide range of techniques to accomplish it. For our purpose, the reconstruction of the complete distribution  $W(x, p)$  is not necessary, since we are only interested in the most efficient way to establish negativity of  $W(x, p)$ . This negativity of  $W(x, p)$  only occurs in limited regions of the phase plane and we thus do not need the full information. This motivates resorting to the so-called filtered back-projection theorem [20, 21], which provides a technique to evaluate the convolution

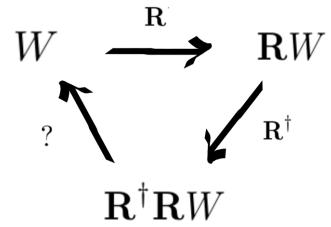


FIG. 5 (colour online). Action of the operators  $\mathbf{R}$  and  $\mathbf{R}^\dagger$ . The inversion cannot be achieved directly through the application of these two operators, *i.e.*, there is a final step missing represented by the question mark. Since the latter is hard to achieve, the idea of the filtered back-projection is to rely on  $\mathbf{R}^\dagger$  for making statements about  $W(x, p)$ .

of  $W(x, p)$  with another function on the phase space. Here, we will only present the main ideas of the theorem and refer to Ref. [20] for details. The Radon transform identifies a linear operator  $\mathbf{R}$  on the functional space of well-behaved functions defined on the phase space, to the space of functions defined on the so-called unit cylinder, spanned by  $(\theta, s)$ . For instance,  $W(x, p)$  belongs to the first functional space while  $\mathbf{R}W(\theta, s)$  belongs to the latter. Both spaces can be equipped with suitable inner products and accordingly, if the Radon transform operator is defined on functions like  $W(x, p)$ , the definition of the adjoint operator  $\mathbf{R}^\dagger$  follows. It can be proven [20, 21] that the action of  $\mathbf{R}^\dagger$  on functions defined on the unit cylinder is given by

$$\mathbf{R}^\dagger \omega(x, p) = \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \omega(\theta, x \cos \theta + p \sin \theta) d\theta. \quad (26)$$

Despite this formula makes the adjoint Radon transformation easy to compute, it does not solve the problem of determining the distribution  $W(x, p)$  because the adjoint  $\mathbf{R}^\dagger$  is not the inverse of  $\mathbf{R}$  (see Fig. 5). Instead of inversion, we can, however, use the adjoint to determine properties of  $W(x, p)$  from the filtered back-projection formula, which states that

$$\begin{aligned} & \int_{\mathbb{R}^2} F(x_0 - x, p_0 - p) W(x, p) dx dp = \\ & = \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \int_{\mathbb{R}} \mathbf{R}W(\theta, x_0 \cos \theta - p_0 \sin \theta - s) \omega(\theta, s) ds d\theta, \end{aligned} \quad (27)$$

provided that the filter,  $F(x, p)$ , and its kernel,  $\omega(\theta, s)$ , are related by the adjoint Radon operator, that is

$$\mathbf{R}^\dagger \omega(x, p) = F(x, p). \quad (28)$$

The advantage of the filtered back-projection formula in Eq. (27) in view of applications is straightforward. By

integrating the right hand side, where  $\mathbf{R}W(\theta, s)$  are the measured distributions and  $\omega(\theta, s)$  is a known function, one attains the convolution of the unknown  $W(x, p)$  and the known filter  $F(x, p)$ . For state reconstruction, the idea is to avoid a direct cumbersome inversion of  $\mathbf{R}$  by using Eq. (27) to get a coarse graining through an *ad hoc* designed filter. It is, however, important to realize that the filtered back-projection theorem shifts the problem of the inversion of the operator  $\mathbf{R}$  to the problem of finding the kernel satisfying Eq. (28) for a given filter  $F(x, p)$ . This problem, also called Cho's problem in the literature, is highly non-trivial. For instance, a solution of Cho's problem for  $F(x, p) = \mathcal{F}(x, p)$  with  $\mathcal{F}$  given in Eq. (3), is not known in general.

For our purpose, the filter  $F(x, p)$ , will play the role of a test function and Eq. (27) gives us a recipe to compute its mean value, possibly translated in the plane, with respect to the unknown distribution  $W(x, p)$ . Hence, similar to the simple test above, the non-existence of a positive probability distribution  $W(x, p)$  can be proven by the right hand side yielding a negative value for a non-negative filter  $F(x, p)$ . The difficulty of solving Cho's problem, however, limits the choice of test functions that we can consider.

## B. Application

In the previous non-classicality test, we looked for a violation of the condition in Eq. (3) for the squeezed and unsqueezed single photon state. Following a similar path of reasoning, we argue that a state is non-classical if it is possible to get a negative outcome by performing the convolution of  $W$  with a non-negative filter function  $F$  (see Eq. (27)) through the back-projection formula with a kernel  $\omega$  as the solution to Cho's problem. In the following, we will refer to the filter  $F$  as the test function. A choice of filter is the normalized characteristic function of a disc, for which Eq. (28) has been solved in the literature [22],

$$F_a(x, p) = \begin{cases} \frac{1}{\pi a^2} & (x^2 + p^2) \leq a^2 \\ 0 & (x^2 + p^2) > a^2 \end{cases} \quad (29)$$

This filter is indeed non-negative and has a compact support so that it can only have a negative expectation value if  $W(x, p)$  takes on negative values. This filter is not usually employed in tomography because it is not robust against noise. Other filters employed in reconstructions are more robust against noise, but they are not non-negative. The convolution with a negative filter would introduce an ambiguity in the interpretation of a possible negative outcome, since one would have to determine if it came from the distribution  $W$  or the filter  $F$ . We, therefore, choose to consider a completely positive filter of the type in Eq. (29) despite its inherent instability. Ideally, it would be better to employ a more smooth non-negative filter, but to our knowledge no simple solutions to Cho's

problem has been determined for such filters. Using a circular disc is, however, only a good test function for the unsqueezed single photon state, which is also rotational invariant. The geometry of the squeezed state (see Fig. 4) suggests that it is better to use a stretched filter  $F_{a,l}(x, p) = F_a(lx, p/l)$  for this state, provided that we can determine the corresponding kernel. Such a stretched filter will have a better overlap with the negative areas in the Wigner function of the squeezed state and we thus expect it to produce statistically more significant results. Note that for the sake of generality, we here allow for a stretching parameter  $l$ , which might be different from the squeezing parameter  $\lambda$ .

The solution to Cho's problem for  $F_a(x, p)$  given by Eq. (29) can be found in Ref. [22] and reads,

$$\omega_a(\theta, s) \equiv w_a(s) = \begin{cases} \frac{1}{\pi a^2} & |s| \leq a \\ \frac{1}{\pi a^2} \left( 1 - \frac{1}{\sqrt{1-(a/s)^2}} \right) & |s| > a \end{cases} \quad (30)$$

In Appendix B, we derive a method to determine the kernel for any linear transformation of the filter. In particular for the stretched filter  $F_{a,l}(x, p)$  we find the corresponding kernel

$$\omega_{a,l}(\theta, s) = \frac{1}{u^2(l, \theta)} \omega_a \left( \frac{s}{u(l, \theta)} \right), \quad (31)$$

where we have defined the function

$$u(\theta, l) = \frac{\cos \theta}{l} \sqrt{1 + l^4 \tan^2 \theta}. \quad (32)$$

In order to use the back-projection formula, we need to determine the integral on the rhs of Eq. (27). Since we know that the Wigner functions in consideration are negative around the origin, we set  $(x_0, p_0) = (0, 0)$  and put  $l = \lambda$ . The motivation for the latter is that if the filter covers the largest possible area in phase space, the error associated to the filtered back-projection is reduced, at least for rather small areas, since we are coarse graining over a larger area of the phase space. On the contrary, a filter limited to a small region around the minimum of  $W(x, p)$  will give a more negative result and thus a larger violation of classicality. The optimal filter is thus obtained as a trade off, which suggests a filter matching the symmetry of  $W(x, p)$ . Consequently, we put  $l = \lambda$ .

We define  $\langle F_{a,\lambda} \rangle$  as being the lhs of Eq. (27), *i.e.*,

$$\langle F_{a,\lambda} \rangle = \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \int_{\mathbb{R}} \mathbf{R}W_\lambda(\theta, s) \omega_{a,\lambda}(\theta, s) ds d\theta. \quad (33)$$

Similar to Eq. (8), this equation expresses the mean value of the positive function  $F$  over the phase space in terms of measurable quantities since, for fixed  $\theta$ , the function  $\mathbf{R}W_\lambda(\theta, s)$  is just the measured distribution of the quadrature  $Q_\theta$ . Nonetheless, the calculation of the mean value of the test function in Eq. (33) involves an integration over both the values of the quadrature and the phases. For the elementary test, only a finite set

of at least  $2N + 1$  phase cuts was required and a negativity was already seen for  $N = 4$  in the single photon state. Here, on the other hand, evaluation of the filtered back-projection over a certain finite number of angles will necessarily involve some ambiguities since it is *a priori* not clear how many angles are required to get a certain precision. Since we do not want to assume *a priori* information about the phase space distribution, a proper analysis ought to include a study of the largest possible error introduced by restricting the number of angles.

As opposed to estimating the error introduced by the finite number of cuts, we will in the following section focus on an alternative strategy. We will refer to this as a Monte Carlo numerical quadrature, where we also consider the measurement angle as a random variable. In this way, the number of cuts is equal to the number of measurements and we thereby gradually remove the ambiguities from the finite number of cuts as the number of measurements increases. We find that this strategy is conceptually simpler to understand, but an alternative and more natural strategy of using a fixed number of phase space cuts might be desirable experimentally. In the main text, we will only discuss this latter strategy briefly and leave most of the discussion to Appendix B where we show that it gives similar results although at the cost of introducing unpleasant ambiguities.

For an experimental test, one is concerned with the statistical error in the evaluation of Eq. (33). This depends on the specific strategy adopted but there is a general issue in all the test we consider, which is related to our choice of kernel  $\omega$ . The discontinuity of the kernel (see Eq. (31)) causes the variance of  $\langle F_{a,\lambda} \rangle$  to be infinite because  $\omega_{a,\lambda}(\theta, s)$  is not square integrable for any  $\lambda$ . The singularity is approached as  $\epsilon^{-1}$  for  $\epsilon \equiv |s - s_0| \rightarrow 0$ , where  $s_0$  is the point where the singularity occurs. This infinity is simply a mathematical ambiguity and has no physical meaning. It can be avoided using a subsidiary kernel

$$\omega_{a,\lambda,\epsilon}(\theta, s) = \begin{cases} 0 & a < |s|/u(\lambda, \theta) < a + \epsilon \\ \omega_{a,\lambda}(\theta, s) & \text{elsewhere,} \end{cases} \quad (34)$$

which corresponds to removing a minimal interval in the values of the quadrature close to the singularity. This results in a small systematic error in the estimation of  $\langle F_{a,\lambda} \rangle$ , equal to

$$\begin{aligned} \Delta_a(\epsilon) &\equiv \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \int_{a u(\lambda, \theta)}^{(a+\epsilon) u(\lambda, \theta)} \frac{\mathbf{R}W_\lambda(\theta, s)}{\pi} |\omega_{a,\lambda}(\theta, s)| d\theta ds \\ &= \int_a^{a+\epsilon} \mathbf{R}W_{\lambda=1}(t) |\omega_{a,\lambda=1}(t)| dt. \end{aligned} \quad (35)$$

Note that this quantity cannot be directly computed from experimental data. Nonetheless, an upper limit on this error can be obtained empirically, by replacing  $\mathbf{R}W$  with its maximum measured value. This leads to an error of the order  $O(\epsilon^{\frac{1}{2}})$ . The variance for the estimate of  $\langle F_{a,\lambda} \rangle$  is thus finite, but proportional to  $|\log \epsilon|$ . In the following,

the use of the the subsidiary kernel  $\omega_{a,\lambda,\epsilon}(\theta, s)$  will always be intended even though we will not explicitly write the subscript  $\epsilon$ .

### C. Monte Carlo numerical quadrature

We now introduce the Monte Carlo numerical quadrature for the double integral appearing in the back-projection formula. The goal is to divide the integrand into two parts,  $\mathbf{R}W_\lambda(\theta, s)\omega_{a,\lambda}(\theta, s)/\pi \equiv p_\lambda(\theta, s) I_{a,\lambda}(\theta, s)$ , where  $p_\lambda(\theta, s)$  is considered to be a probability distribution for two random variables  $\theta$  and  $s$ , and  $I_{a,\lambda}(\theta, s)$  is a suitable integrand reminder. Note that in the experiment, only the variable  $s$ , which characterises the quadrature value, is a truly random variable resulting from the probabilistic nature of the measurement, whereas the phase angle can be controlled experimentally. Nonetheless, we assume that the phase angle is also varied randomly so that, for each shot of the experiment, the experimentalist picks the phase angle according some probability distribution  $\mathcal{C}(\theta)$ . According to Eq. (33), we can then express the filtered back projection as

$$\langle F_{a,\lambda} \rangle = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \int_{\mathbb{R}} I_{a,\lambda}(\theta, s) p_\lambda(\theta, s) ds d\theta. \quad (36)$$

with

$$\begin{cases} p_\lambda(\theta, s) \equiv \mathbf{R}W_\lambda(\theta, s) \mathcal{C}(\theta) \\ I_{a,\lambda}(\theta, s) \equiv \frac{\omega_{a,\lambda}(\theta, s)}{\pi \mathcal{C}(\theta)}, \end{cases} \quad (37)$$

where  $\mathcal{C}(\theta)$  is any distribution of cuts normalized to unity, *i.e.*,  $\int \mathcal{C}(\theta) d\theta = 1$ . At this point, all distributions of cuts are logically equivalent and will give the same result for  $\langle F_{a,\lambda} \rangle$ , but as we discuss below the choice of distribution will make a difference for the efficiency of the numerical quadrature.

In an experiment, we assume to obtain a set of  $M$  outcomes of the form  $\{(\theta_i, s_i)\}$ , distributed according to  $p_\lambda(\theta, s)$ , so that the average of the test function can be estimated as

$$\langle F_{a,\lambda} \rangle \simeq \frac{1}{M} \sum_{i=1}^M I_{a,\lambda,\epsilon}(\theta_i, s_i). \quad (38)$$

The estimate can then be used to check the classicality condition in Eq. (3). For the statistical significance of the test we, however, need to include the error in the Monte Carlo quadrature. The variance of the estimate  $\langle F_{a,\lambda} \rangle$  can be expressed as

$$\sigma_{F_{a,\lambda}}^2 = \frac{1}{(M-1)} \sum_{i=1}^M \left( I_{a,\lambda}(\theta_i, s_i) - \langle F_{a,\lambda} \rangle \right)^2 \quad (39)$$

Accordingly, the relevant quantity to consider for the violation is the ratio

$$\frac{\langle F_{a,\lambda} \rangle}{\sigma_{F_{a,\lambda}} + \Delta_a(\epsilon)}, \quad (40)$$

which accounts both for the statistical uncertainty  $\sigma_{F_{a,\lambda}}$  and the systematic error  $\Delta_a(\epsilon)$ . Since we now consider not only the mean but also the statistical uncertainty, such a ratio depends on the distribution of cuts  $\mathcal{C}(\theta)$ , which plays an essential role in the minimisation of  $\sigma_{F_{a,\lambda}}$ . This is discussed below and in Appendix B, where we consider both the squeezed and unsqueezed single photon state. Now there is no longer a sensible way to remove the dependence on the number of measurements  $M$  from this ratio. In fact, the asymptotic behaviour for large  $M$  is no longer  $\sim \sqrt{M}$  as for the simple test because there is a trade off between  $\sigma_{F_{a,\lambda}}$ , diverging as  $|\log \epsilon|$ , and  $\Delta_a(\epsilon)$  decreasing as  $\epsilon^{\frac{1}{2}}$ . Therefore it is not convenient to rescale it by  $\sqrt{M}$  and take the limit of  $M \rightarrow \infty$  as we did for the simple test (see Eq. (14)). In order to achieve a comparison with the simple test discussed above we consider the ratio

$$\mathcal{R}(M) \equiv \frac{\langle \mathcal{F} \rangle}{\sigma_{\mathcal{F}}} \frac{\sigma_{F_{a,\lambda}} + \Delta_a(\epsilon)}{\langle F_{a,\lambda} \rangle}, \quad (41)$$

which describes the ratio between the statistical certainty of the tests, *i.e.*, the ratio of the number of standard deviations with which the tests violate the classicality criterion. A value  $\mathcal{R}(M) > 1$  ( $\mathcal{R}(M) < 1$ ) thus signifies that the simple test (inverse Radon method) has a larger statistical certainty.

#### D. Single photon state.

First, we consider the back-projection formula for the unsqueezed single photon state. The rotational invariance of the Wigner function  $W(x, p)$  motivates a phase cut distribution of the form

$$\mathcal{C}(\theta) \equiv \frac{1}{\pi}, \quad (42)$$

which corresponds to a uniform, flat distribution of the phase cuts. As shown in Appendix B, this is the optimal distribution of cuts for minimizing  $\sigma_{F_{a,\lambda}}$ . Consequently, we have  $I(\theta, s) \equiv \pi \omega_{a,\lambda}(\theta, s)$ .

For given  $M$ , we have minimized the value of  $\frac{\langle F_{a,\lambda} \rangle}{\sigma_{F_{a,\lambda}} + \Delta_a(\epsilon)}$  with respect to the remaining free parameters  $a$  and  $\epsilon$ , keeping  $\lambda = 1$  because of the rotational symmetry. In the minimisation, we numerically compute the double integral that estimate the value of the mean  $\langle F_{a,\lambda} \rangle$  and variance  $\sigma_{F_{a,\lambda}}$  for given  $a$  and  $\epsilon$  and analytically compute the upper bound of the systematic error  $\Delta_a(\epsilon)$ . Consequently, we can evaluate the ratio in Eq. (40) and run a numerical optimisation to minimize it. We compare the result of the minimisation to the best elementary test we could achieve with  $N = 16$  ( $N$  is the polynomial degree) by calculating the ratio  $\mathcal{R}(M)$ . The results of this is shown in Fig. 6 as a function of  $M$ . From the optimisation we find that the systematic error decreases as a power law and this causes the variance of the kernel to grow logarithmically. Consequently, the

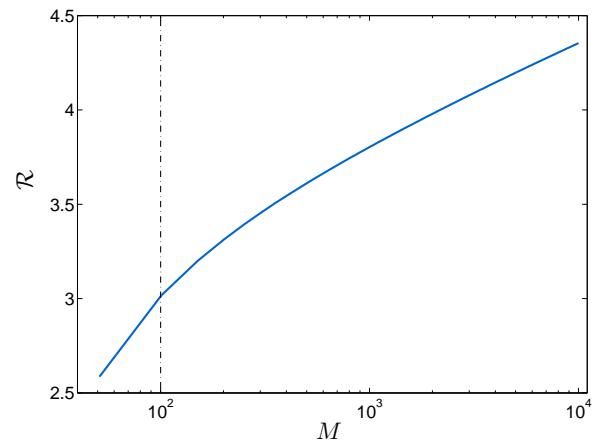


FIG. 6 (colour online). Plot of the ratio  $\mathcal{R}(M)$  of the statistical significance of the elementary test and the inverse Radon method for the single photon state. The comparison is with respect to the best elementary test for  $N = 16, m = 17$ . The value  $\mathcal{R}(M) > 1$  indicates that the elementary test is statistically more reliable. The almost linear part of the graph shows a nearly logarithmic growth for high  $M$ ; the low  $M$  region deviates from this behaviour due to the  $\sqrt{M - m}$  dependence in  $\sigma_{\mathcal{F}}$ . The dashed vertical line indicate the threshold for the inverse Radon method to violate the non-classicality by two standard deviations.

quantity  $\mathcal{R}(M)$  does not settle to an asymptotic value, but keeps increasing as a function of  $M$  because of the trade off between the systematic error and variance of the kernel. For all values of  $M$  we find  $R > 1$  which means that the elementary test is statistically more significant. The inverse Radon is closer to the elementary test for a small number of measurements. Note that the very small  $M$  region cannot be investigated since  $\sigma_{\mathcal{F}}$  is defined only for  $M > m$ , the least number of cuts being  $m = 17$  for  $N = 16$ . This region is, however, not very interesting since it lies in the region where the inverse Radon is not capable of producing a violation by two standard deviations (dashed vertical line in Fig. 6).

#### E. Squeezed state.

We now consider the squeezed single photon state, where an explicit optimization in the choice of  $\mathcal{C}(\theta)$  is crucial for obtaining an effective test. As shown in Appendix B, the naive approach of choosing  $\mathcal{C}(\theta)$  as given by Eq. (42) leads to a poor scaling of the variance as  $\sim \lambda^2$  ( $\sim \lambda^{-2}$ ) for  $\lambda \gg 1$  ( $\lambda \ll 1$ ). This result shows that for the squeezed state a uniform distribution of phase cuts is not desirable since the squeezed state has a highly non-uniform angular distribution. It is thus better to have more measurements in regions of phase space with a rapid variation. To find a more optimal distribution  $\mathcal{C}(\theta)$ , we use the following result from the theory of Radon trans-

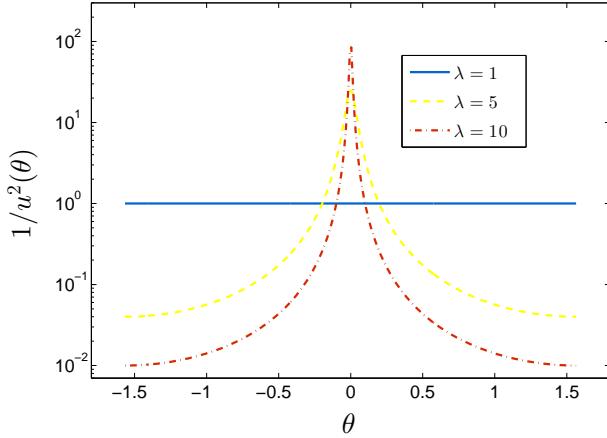


FIG. 7 (colour online). Logarithmic plot of  $\frac{1}{u^2(\theta, \lambda)}$  as a function of  $\theta$  for some values of the squeezing parameter  $\lambda$ . The region that contributes the most to the integral in Eq. (44) gets narrower as lambda increases.

form [20]

$$\mathbf{RW}_\lambda(\theta, s) = \frac{1}{u(\lambda, \theta)} \mathbf{RW} \left( \frac{s}{u(\lambda, \theta)} \right) \quad (43)$$

together with the transformed kernel for the stretched filter given in Eq. (31). Using these two results, we find that

$$\begin{aligned} \langle F_{a,\lambda} \rangle &= \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \int_{\mathbb{R}} \mathbf{RW}_\lambda(\theta, s) \frac{\omega_{a,\lambda}(\theta, s)}{\pi} ds d\theta \\ &= \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{1}{u^2(\lambda, \theta)} \int_{\mathbb{R}} \mathbf{RW}_{\lambda=1}(t) \frac{\omega_{a,\lambda=1}(t)}{\pi} dt d\theta. \end{aligned} \quad (44)$$

Eq. (44) shows that not all angles contribute equally to the final result. Because of the use of a stretched filter most of the weight is focused on a small subset of angles. This is illustrated in Fig. 7 where we plot  $\frac{1}{u^2(\theta, \lambda)}$  as a function of the phase  $\theta$  for various values of  $\lambda > 1$ . As shown in the figure,  $\frac{1}{u^2(\theta, \lambda)}$  gives more weight to phases close to  $\theta = 0$  ( $\theta = \pi$ ) for  $\lambda > 1$  ( $\lambda < 1$ ). Since it is important to get an accurate sampling of the most relevant part of the integral, it will be advantageous to choose a probability distribution for the phase cuts  $\mathcal{C}(\theta)$  which is higher around this region. As shown in Appendix B, the optimal distribution of cuts is indeed

$$\mathcal{C}(\theta) = \frac{1}{\pi u^2(\lambda, \theta)}, \quad (45)$$

which coincides with Eq. (42) for  $\lambda = 1$ . Furthermore, we show in the appendix that the relevant ratio  $\frac{\langle F_{a,\lambda} \rangle}{\sigma_{F_{a,\lambda}} + \Delta_a(\epsilon)}$  does not depend on  $\lambda$  for this choice of phase cut distribution, *i.e.*, it is thus equal to the one found for the rotationally symmetric state. This is completely equivalent to the analogous result found for the simple test.

Hence the same conclusions drawn about the statistical significance of the elementary and inverse Radon tests for the single photon state also applies to the squeezed single photon state.

## F. Finite Number of Cuts

In experiments, a strategy more widely used than the Monte Carlo quadrature is to have a small set of cuts but many measurements for each cut. In order to study this situation, it is more convenient to work out the right hand side of Eq. (27) in a different way than for the Monte Carlo quadrature. This is done in two steps. First, we introduce the function

$$g_{a,\lambda}(\theta) \equiv \int_{\mathbb{R}} \mathbf{RW}_\lambda(\theta, s) \omega_{a,\lambda}(\theta, s) ds, \quad (46)$$

which is assumed to be accurately estimated by having many measurements per cut. The integral over the phase cuts is then written as a summation over a convenient choice of cuts  $\{\theta_j\}$  with weights  $w_j$

$$\frac{1}{\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} g_{a,\lambda}(\theta) d\theta \rightarrow \sum_{\{\theta_j\}} w_j g_{a,\lambda}(\theta_j). \quad (47)$$

The discretisation of the integral above, formally called numerical quadrature, is the crucial step in a strategy with a finite number of cuts. This choice, however, gives rise to ambiguities because we cannot completely know the function  $g_{a,\lambda}(\theta)$  since it is sampled with only a few phase arguments  $\{\theta_j\}$ . The first ambiguity is in the choice of the number of phase cuts. We cannot *a priori* determine a minimum number of phase cuts required as in the elementary test. Regardless of the number of cuts, the full expression on the left hand side of Eq. (47) involves an integral. Hence one can always wonder if adding more phase cuts could reveal narrow features missed by the finite number of cuts. It is therefore in principle necessary to constrain the influence of such narrow features on the final result to convincingly demonstrate non-classicality. Another ambiguity is in the estimation of the error due to the numerical quadrature, which cannot be easily determined unless some other information about  $g_{a,\lambda}(\theta)$  is known.

Despite of these ambiguities, it is still reasonable to ask whether we can determine an efficient numerical quadrature that uses a low number of phase cuts compared to the total number of measurements. In other words, we want to understand whether the inverse Radon method can provide a way to choose a suitable small set of phase cuts if some further sensible assumptions are made. To give an answer to this question, in the analysis carried out in Appendix B, we assume *ab initio* that the states considered are the single photon state, described by a symmetric  $W_\lambda(x, p)$ , as well as the squeezed single photon state. For our theoretical analysis, such an hypothesis

relies ultimately on the Wigner function to be the correct description of the state, which is somewhat logically problematic since this information is exploited to carry out the method, not just to optimise it.

We show in Appendix B that this strategy can also be carried out in such a way that the statistical strength of the test is the same for both rotationally symmetric and squeezed states. As the squeezing increases, however, more cuts are needed in order to keep the uncertainty of the test below a certain threshold and these cuts should be highly asymmetrically distributed. Choosing such an asymmetric distribution can be justified by the asymmetry of the filter function, which is highly peaked around a narrow distribution of angles. On the other hand, it is hard to quantitatively account for the absence of further errors due to the unknown angular dependence of the underlying physical state, and this introduces an ambiguity in the test.

In essence, closing such “loop-holes” in the argument is what is achieved with the elementary test, which is explicitly designed to deal with a finite number of cuts. Due to the ambiguities discussed here, making a more direct comparison between the elementary test and a strategy with a finite number of cuts is highly non-trivial and is thus beyond the scope of this article.

#### IV. CONCLUSION AND DISCUSSION

In conclusion, we have generalised the non-classicality test presented in Ref. [6] and developed a framework for finding the optimal test function for a given quantum state. In the generalisation of the test, we consider test functions which are arbitrary polynomials in the phase space. Thereby, we allow test functions which do not have rotational symmetry and are therefore more suited for testing the non-classicality of quantum states which do not possess this symmetry. As a specific example, we have found the optimal test function for a squeezed single photon state and shown that it results from a simple transformation of the rotationally invariant test function considered in Ref. [6].

Furthermore, we have performed a comparison between the simple non-classicality test and more conventional tests based on the inverse Radon transform. In order for the comparison to be fair, we have extended the strategy of inverse Radon transformation based on the filtered back-projection formula so that it is more optimised for squeezed states. The constraints of the filtered back-projection formula and the limited number of known completely positive filter functions meant, however, that only a rather restricted class of test functions could be investigated (as compared with the elementary test, where there is a lot of freedom). In addition, the filtered back-projection formula needs to be handled with care because it introduces problems with divergencies and ambiguities from introducing a finite number of cuts. With a Monte Carlo quadrature method the ambiguities

from the finite number of cuts can be avoided, but this requires the use of as many quadrature cuts as the number of measurements. Reducing the number of cuts introduces undesired ambiguities, which may question the whole outcome of the test. In particular, these ambiguities are severe for strongly squeezed states where it is necessary to have a large number of cuts in a small angular interval where the integrand varies rapidly. Without any prior information about the state, such a choice of distribution introduces ambiguities, which make it hard to give a convincing proof of the non-classicality of a state. All of these problems are avoided with the elementary test. In addition to this, we find that the elementary test is statistically stronger than the methods based on the inverse Radon transform (at least compared to the version based on the Monte Carlo quadrature method).

We thus find that the simple test is more suited and straightforward than the inverse Radon transform for experimental demonstration of non-classicality. The advantage of the simple test originates from the fact that it is naturally designed as a non-classicality test under the constraints imposed in a typical experiment. In particular, the elementary test is specifically designed to work with a finite number of quadratures being measured. On the contrary, the inverse Radon transform is a more general method, which allows to gain more information about the physical state but it is less suited for this specific situation. The statistical weakness of the latter could, however, be due to the fact that the kernels considered here are not ideal due to their discontinuity. On the other hand, the smooth filters considered in the literature that grant a better convergence are not necessarily completely positive. A negativity observed with such a filter would therefore inevitably introduce the question of whether the result was caused by the negativity of the filter and this could compromise the test.

We have considered the simple test for a quadrature squeezed single photon state but the general framework developed here can be applied to any quantum state or any physical system. In particular, it could be interesting to use the test to demonstrate non-classicality of opto-mechanical systems [10, 23]. Currently much effort is devoted to reaching the quantum regime for these systems. To convincingly show that such systems are outside the realm of classical physics, it is essential to demonstrate a contradiction with classical physics along the lines discussed here. Also it could be interesting to apply the method to investigate, the negativity of macroscopic quantum states such as large Schrödinger cat states and potentially use this as a measure of non-classicality in such systems. Finally, one could imagine using the elementary test method as a general tool for tomography without resorting to the heavy machinery of inverse Radon transformations. Since we find the method to be more reliable and less ambiguous, one could imagine other situations where this method could be a desirable strategy for tomography.

## ACKNOWLEDGMENTS

We thank Eran Kot for valuable discussions and gratefully acknowledge financial support from the Lundbeck foundation, the Carlsberg foundation and the Famiglia Legnanese foundation. The research leading to these results was funded by the European Union Seventh Framework Programme through ERC Grant QIOS (Grant No. 306576).

## Appendix A: Test function method

In this appendix, we present the details of the results on the simple test method discussed in the main text. First, we prove, in three intermediate steps, a slightly more general version of Theorem 1. The first step is to establish that for a rotationally symmetric distribution the average value of any test function can be expressed as the average value of a radial test function. The second step is to identify the tests that minimise the variance for a given value of the mean. Finally, the explicit construction of a radial test which minimises the variance is provided. The proof of the theorem is then obtained by combining these three partial results. Second, we prove Theorem 2 concerning squeezed states.

### 1. General formulation.

We will take into account a slightly more general form of the test function than in the main text in order to derive the proofs of the theorems. To begin, we consider the test function to have the form

$$\mathcal{F}(x, p) = 1 + \sum_{i=1}^{2N} P_i(x, p) \quad (\text{A1})$$

where  $P_i(x, p)$  is a homogeneous polynomial of degree  $i$  defined on the phase space of  $x$  and  $p$ . Moreover, all the test functions of the form in Eq. (A1) are implicitly taken to be non-negative, thus we are more general than in the main text where we only considered square polynomials. Note that the polynomials  $P_i$  can be expressed in terms of the functions  $x^{i-k} p^k$  to recover the notation employed in the main text of the paper. By assumption, we expect  $\mathcal{F}$  to be in the algebra of observables, formally defined as the commutative algebra generated by all the quadratures. In particular, for fixed  $i$ , any  $P_i(x, p)$  belongs to a linear subspace of this algebra, together with  $Q_\theta^i$  for any  $\theta$ . Any set of at least  $(i+1)$  functions of the form  $Q_\theta^i$  for different  $\theta$  is a (possibly overcomplete) set in this linear subspace. Accordingly, given a set of  $j = 1, \dots, m$  different cuts  $\{\theta_j\}$ , with  $m \geq 2N + 1$ , all the polynomials  $P_i(x, p)$  involved in Eq. (A1) can be expressed as

$$P_i(x, p) = \sum_{j=1}^{2N+1} T_{i;j} Q_{\theta_j}^i, \quad (\text{A2})$$

with  $T_{i;j}$  being real-valued but not uniquely defined. We will express a test,  $\mathbb{T}$  as

$$\mathbb{T} \equiv (N, P_i, \theta_j, T_{i;j}), \quad (\text{A3})$$

where the non-negativity of the test function is intended, but not explicitly stated. This is the most general class of tests we consider and we refer to this class as  $\mathcal{C}$ . A test of the form in Eq. (A3) belongs to this class,  $\mathbb{T} \in \mathcal{C}$ , if  $\mathcal{F}$  as defined in Eq. (A1) is non-negative and Eq. (A2) holds.

The rotated coordinates,  $Q_\phi$  and  $P_\phi$  are defined as

$$\begin{pmatrix} Q_\phi \\ P_\phi \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix} \quad (\text{A4})$$

and correspond to the quadrature  $Q_\phi$  and the rotated canonical momentum  $P_\phi$ . Accordingly, we define the rotation of polynomials by

$$P_i[\phi](x, p) \equiv P_i(Q_\phi, P_\phi). \quad (\text{A5})$$

We refer to a test as a radial test and denote it by  $\mathbb{T}_{rad}$ , if  $P_i[\phi](x, p) = P_i(x, p) \quad \forall \phi \in \mathbb{R}$ , for all  $i = 1, \dots, 2N + 1$ . Consequently, in this case, the test function,  $\mathcal{F}$ , is only a function of the radius. The class of such radial tests is denoted  $\mathcal{C}_{rad}$  and it is obviously a proper subclass of  $\mathcal{C}$ . To complete the set of notations, it is sensible to explicitly express the dependence of  $\langle \mathcal{F} \rangle / \sigma_{\mathcal{F}}$  in terms of the free parameters of the test. Therefore we shall write

$$\frac{\langle \mathcal{F} \rangle}{\sigma_{\mathcal{F}}} \equiv \mathcal{G}[\mathbb{T}, M], \quad (\text{A6})$$

so that the actual figure of merit  $\mathcal{G}$  is simply recovered as  $\mathcal{G}[\mathbb{T}] = \lim_{M \rightarrow \infty} \frac{\mathcal{G}[\mathbb{T}, M]}{\sqrt{M-m}}$

### 2. Rotationally invariant states.

As hinted above, even though our objective is the optimal  $\mathcal{G}[\mathbb{T}, M]$  for any  $M$ , we do not yet consider the direct optimisation with respect to the free parameters of the test. Instead, we first want to understand the structure of the class  $\mathcal{C}$  and its relation to  $\mathcal{C}_{rad}$  for a rotationally symmetric  $W(x, p)$ . The first result concerns the mean value  $\langle \mathcal{F} \rangle$  and, in particular how the class  $\mathcal{C}$  and  $\mathcal{C}_{rad}$  are related under the condition that all  $\langle P_i \rangle$  are given.

**Proposition A.1.** *Assume that  $W(x, p)$  is rotationally symmetric and consider a generic test  $\mathbb{T} \in \mathcal{C}$  such that  $\mathcal{F} = 1 + \sum_{i=1}^{2N} P_i(x, p)$  is the test function. There always exists a radial test  $\mathbb{T}_{rad} \in \mathcal{C}_{rad}$  with test function of the form  $\mathcal{F}_{rad} = 1 + \sum_{i=1}^{2N} \bar{P}_i(x, p)$ , where  $\bar{P}_i(x, p)$  is rotationally symmetric and fulfils  $\langle \bar{P}_i \rangle = \langle P_i \rangle$  for any  $i$ , so that  $\langle \mathcal{F} \rangle = \langle \mathcal{F}_{rad} \rangle$ . Furthermore, all the radial tests that satisfy this condition differ from each other only in the parameters  $\{\theta_j\}$ ,  $\{T_{i;j}\}$ .*

*Proof.* The proof is by construction. The given test is characterized by  $\mathbb{T} \equiv (N, P_i, \theta_j, T_{i;j})$ , and we define

$$\bar{P}_i \equiv \frac{1}{2\pi} \int_0^{2\pi} P_i[\theta] d\theta \quad \forall i = 1, \dots, 2N. \quad (\text{A7})$$

Note that  $\bar{P}_i$  is rotationally invariant since

$$\bar{P}_i[\phi] = \frac{1}{2\pi} \int_0^{2\pi} P_i[\theta + \phi] d\theta = \bar{P}_i. \quad (\text{A8})$$

Moreover, because of the rotational symmetry of  $W(x, p)$ , a simple change of variables to the rotated coordinates shows that  $\langle P_i[\theta] \rangle = \langle P_i \rangle$  for any  $\theta$ . Thus,

$$\begin{aligned} \langle \bar{P}_i \rangle &= \frac{1}{2\pi} \int_0^{2\pi} \langle P_i[\theta] \rangle d\theta \\ &= \langle P_i \rangle \frac{1}{2\pi} \int_0^{2\pi} d\theta = \langle P_i \rangle. \end{aligned} \quad (\text{A9})$$

Considering  $\mathcal{F}_{rad} = 1 + \sum_{i=1}^{2N} \bar{P}_i(x, p)$ , it can be seen that the rotationally invariant test function is non-negative since

$$\begin{aligned} \mathcal{F}_{rad}(x, p) &= 1 + \sum_{i=1}^{2N} \frac{1}{2\pi} \int_0^{2\pi} P_i[\theta](x, p) d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \mathcal{F}[\theta](x, p) d\theta, \end{aligned} \quad (\text{A10})$$

i.e., for fixed  $(x, p)$ , it is the integral of a non-negative function  $\mathcal{F}_{rad}[\theta](x, p)$  wrt. the variable  $\theta$ . The statement of Proposition A.1 follows immediately. Note that no consideration of  $\theta_j$  or  $T_{i;j}$  was necessary because they are irrelevant for the mean value of  $\mathcal{F}$ .  $\square$

Above, we have proven that for a rotationally symmetric distribution,  $W(x, p)$ , and a non-rotational test, we can always find a rotationally symmetric version of the test, which gives the same average. We will now show that the rotationally invariant test will also have a smaller uncertainty in the mean value. To do so, we initially study a more general problem, neglecting for now the non-negative property of the test functions. Since we consider a more general class of functions, we will at first (see Proposition A.2) achieve results without considering the requirements for proper test functions, and thereby of no direct physical significance. Later we will, however, show (see Proposition A.3) that it is indeed possible to find a physically meaningful test fulfilling the conditions of a optimum in the broader class, and thus also an optimum in the restricted class.

First, we give a suitable expression for  $\sigma_{\mathcal{F}}^2$  that accounts for the rotational symmetry of  $W(x, p)$ . Since  $\langle Q_{\theta}^i \rangle \equiv \langle Q^i \rangle$  does not depend on  $\theta$  for any  $i$ , we shall introduce the covariance matrix of the quadratures,

$$\frac{1}{2} g^{ii'} \equiv \langle Q^{i+i'} \rangle - \langle Q^i \rangle \langle Q^{i'} \rangle \quad (\text{A11})$$

It can then be seen that  $\sigma_{\mathcal{F}}^2 \equiv \langle \mathcal{H}_j^2 \rangle - \langle \mathcal{H}_j \rangle^2 = \frac{1}{2} \sum_{i,i'=1}^{2N} g^{ii'} T_{i;j} T_{i';j}$  so that a final expression for  $\sigma_{\mathcal{F}}^2$  is achieved

$$\sigma_{\mathcal{F}}^2 = \frac{1}{2} \sum_{j=1}^m \sum_{i,i'=1}^{2N} g^{ii'} \frac{T_{i;j} T_{i';j}}{M_j - 1}, \quad (\text{A12})$$

where  $\{M_j\}$  is the number of measurements for the  $j$ 'th cut and  $\sum_j M_j = M$ .

We now perform an optimisation in the class of functions of the form  $\mathcal{F} = 1 + \sum_{i,j} T_{i;j} Q_{\theta_j}^i$ , with the only constraint  $\sum_j T_{i;j} = K_i$ . This constraint on  $T_{i;j}$  originates from the identity in Eq. (A2). In the case of rotational symmetry, where  $\langle Q_{\theta_j}^i \rangle \equiv \langle Q^i \rangle$  does not depend on  $\theta_j$ , this implies that

$$\sum_{j=1}^m T_{i;j} = \frac{\langle P_i \rangle}{\langle Q^i \rangle} \equiv K_i, \quad (\text{A13})$$

where  $K_i$  are constants that depend only on the values  $\langle P_i \rangle$  and no other free parameters of the test. By optimising the variance subject to this constraint, we thus optimise over the class of tests giving a fixed value of the mean value. Since the true optimum must also be the optimum for the particular mean value obtained, properties proved under this constraint also applies to the true optimum. The optimisation is done keeping  $N, M$  and  $m$  fixed, since these parameters do not depend on whether the test is radial or general. We can thus state and prove the following result.

**Proposition A.2.** *Assume  $\mathcal{F}$  to be any generic polynomial function of order  $2N$  over an arbitrary set of  $m$  quadratures, that is*

$$\mathcal{F} = 1 + \sum_{j=1}^m \sum_{i=1}^{2N} T_{i;j} Q_{\theta_j}^i \quad (\text{A14})$$

*We assume that all the quadratures  $Q_{\theta_j}$  are independent, identically distributed random variables and that  $T_{i;j}$  are only constrained by the relation  $\sum_{j=1}^m T_{i;j} = K_i$ . It follows that the minimum value of  $\sigma_{\mathcal{F}}^2$ , after optimising  $T_{i;j}$  and keeping fixed  $N, M$  and  $m$ , is*

$$\sigma_{\mathcal{F}}^2 = \frac{1}{2} \frac{\sum_{i,i'=1}^{2N} g^{ii'} K_i K_{i'}}{M - m} \quad (\text{A15})$$

*and is thus independent of the particular choice of  $\{M_j\}$  satisfying  $\sum_j M_j = M$ . Furthermore, if  $\{M_j\}$  are all the same, the best  $T_{i;j}$  is independent of  $j$ , i.e.,  $T_{i;j} = K_i/m$ ,  $m$  being the number of cuts.*

*Proof.* Consider a distribution of measurements  $\{M_j\}$ . For simplicity, we define  $\tilde{M}_j = M_j - 1$  so that  $\sum_j \tilde{M}_j = M - m$ . The constrained optimisation problem is then

studied using the Lagrangian function

$$\begin{aligned} \mathcal{L}(\{T_{i;j}\}; \{\lambda^i\}) &= \sigma_{\mathcal{F}}^2 - \sum_{i=1}^{2N} \lambda^i \sum_{j=1}^m T_{i;j} \\ &= \sum_{j=1}^m \left\{ \frac{1}{2} \sum_{i,i'=1}^{2N} \frac{g^{ii'} T_{i;j} T_{i';j}}{\tilde{M}_j} - \sum_{i=1}^{2N} \lambda^i T_{i;j} \right\}, \end{aligned} \quad (\text{A16})$$

which only depends parametrically on  $\{\tilde{M}_j\}$ . The optimal solution is determined by

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial T_{i;j}} = \sum_{i'=1}^{2N} \frac{g^{ii'} T_{i';j}}{\tilde{M}_j} - \lambda_i = 0 & \forall i, j \\ \sum_{j=1}^m T_{i;j} - K_i = 0 & \forall i \\ \sum_{j=1}^m \tilde{M}_j = M - m. \end{cases} \quad (\text{A17})$$

It is seen from the first equation that the optimal solution is such that  $T_{i;j}/\tilde{M}_j$  does not depend on  $j$ , i.e., it is independent of the choice of phase cuts. From this, we find that the solution fulfilling all three conditions is

$$T_{i;j} = \frac{K_i \tilde{M}_j}{M - m}. \quad (\text{A18})$$

Then, the constrained stationary value of the function in Eq. (A16) is

$$\sigma_{\mathcal{F}}^2 = \frac{1}{2} \frac{\sum_{i,i'=1}^{2N} g^{ii'} K_i K_{i'}}{M - m}. \quad (\text{A19})$$

It can be proven that it is a minimum by noting that the Hessian of Eq. (A15) is the tensor product of the covariance matrix and the diagonal matrix containing  $1/\tilde{M}_j$  on the diagonal. Since both of these matrices are positive-semidefinite by definition the found extremum is a minimum. Finally, it follows from Eq. (A18) that when  $\{M_j\}$  are all the same, the ratio  $\frac{\tilde{M}_j}{M - m}$  is exactly  $1/m$ .  $\square$

The crucial part of the theorem is the independence of the covariance matrix in Eq. (A11) on the choice of phase cuts, which is a direct consequence of the rotational symmetry of  $W(x, p)$ . It follows that the choice of the cuts is irrelevant for Proposition A.2. Furthermore, the minimum of  $\sigma_{\mathcal{F}}^2$  does not depend on the choice of  $\{M_j\}$ , provided that  $M$  is fixed. This is in contradiction with the physical intuition that it is not a suitable strategy to perform most of the measurements on a single quadrature whereas only a few measurements are performed on the remaining quadratures. Note, however, that this result above is obtained in the broader class of functions only subject to the constraint  $\sum_j T_{i;j} = K_i$ . Hence, for arbitrary choice of  $\{M_j\}$ , we might not be able to construct radial polynomials fulfilling the general requirement Eq. (A2) with the optimal  $T_{i;j}$  given by Proposition A.2. As we will show below, it is, however, possible to find a proper test fulfilling both Eq. (A2) and

the optimum in Proposition A.2. To this end, it is sensible to fix  $M_j$  to be all the same, so that the statistics per cut is unbiased. Then we will show that we can construct radial polynomials according to Eq. (A2) in such a way that  $T_{i;j}$  is independent of  $j$ . Note that the only radial polynomial we can construct must be functions of  $(x^2 + p^2)$  to retain rotational symmetry.

**Proposition A.3.** *Assume  $T_{i;j}$  to be independent of  $j$ . Then the polynomials  $P_i$  given by the identity in Eq. (A2) are radial polynomials provided that the phase cuts are uniformly distributed. In particular, the identity*

$$(x^2 + p^2)^n = T_{2n} \sum_{j=1}^m Q_{\theta_j}^{2n} \quad (\text{A20})$$

holds, where  $T_{2n} = \binom{2n}{n}^{-1} \frac{4^n}{m}$  and  $\theta_j = \frac{j\pi}{m}$  with  $m \geq n + 1$ .

Note that we have dropped the  $j$  index in the coefficient  $T_{i;j} \Big|_{i=2n} \equiv T_{2n}$  and we have used the index  $n$  in place of  $i$  to avoid confusion in the subsequent proof of the statement, where the imaginary unit is used.

*Proof.* Since  $T_{i;j}$  do not depend on  $j$ , we consider the following sum

$$\sum_{j=1}^m Q_{\theta_j}^{2n} \quad (\text{A21})$$

where for the sake of generality, the sum is extended over  $m$  cuts, with the suitable  $m$  to be determined. Using the exponential representation of the sine and cosine functions, we get

$$\begin{aligned} Q_{\theta_j}^{2n} &= \frac{1}{4^n} (e^{i\theta_j} (x - ip) + e^{-i\theta_j} (x + ip))^{2n} \\ &= \frac{1}{4^n} \sum_{l=0}^{2n} \binom{2n}{l} (x - ip)^{2n-l} (x + ip)^l e^{i2\theta_j(n-l)} \end{aligned} \quad (\text{A22})$$

so that

$$\begin{aligned} \sum_{j=1}^m Q_{\theta_j}^{2n} &= \\ &= \frac{1}{4^n} \sum_{l=0}^{2n} \binom{2n}{l} (x - ip)^{2n-l} (x + ip)^l \left( \sum_{j=1}^m e^{i2\theta_j(n-l)} \right). \end{aligned} \quad (\text{A23})$$

By choosing uniformly distributed cuts, we can turn the summation over  $j$  into a simple geometric series. Letting  $\theta_j = \frac{j\pi}{m}$ , the geometric series is

$$\sum_{j=1}^m \left( e^{i \frac{(n-l)}{m} 2\pi} \right)^j \quad (\text{A24})$$

and the exponential is different from unity if  $\frac{n-l}{m}$  is not an integer. Accordingly, if  $m \geq n+1$ , we have that

$$\sum_{j=1}^m e^{i2\theta_j(n-l)} = m\delta_{l,n} \quad (\text{A25})$$

Inserting this into Eq. (A23) gives

$$\sum_{j=1}^m Q_{\theta_j}^{2n} = \frac{m}{4^n} \binom{2n}{n} (x^2 + p^2)^n \quad (\text{A26})$$

and hence the statement of the proposition is proven.  $\square$

Note that only  $n+1$  cuts are needed to express a polynomial of degree  $2n$  such as  $(x^2 + p^2)^n$ . This is because odd powers of  $x$  and  $p$  are not involved. Therefore, for a radial test function of degree  $2N$  (corresponding to the square of a polynomium of degree  $N$ ), we only need a minimum number of  $N+1$  cuts.

We now have all the intermediate results necessary to prove the statement of Theorem A.1, which constitutes the main result concerning the optimal  $\mathcal{G}[\mathbb{T}, M]$ .

**Theorem A.1.** *Assume that  $W(x, p)$  is rotationally symmetric and takes on negative values; then the optimal test, i.e., the test  $\mathbb{T}_{\text{best}}$  that minimises  $\mathcal{G}[\mathbb{T}, M]$  for any  $M$  belongs to the subclass  $\mathcal{C}_{\text{rad}}$  and the cuts are uniformly distributed on the plane (see Eq. (A20)).*

*Proof.* Let us fix  $M$  and pick up a generic test  $\mathbb{T} \in \mathcal{C}$  characterized by  $(N, P_i, \theta_j, T_{i;j})$ , for which  $\mathcal{G}[\mathbb{T}, M]$  is negative. We will construct a radial test  $\mathbb{T}_{\text{rad}} \in \mathcal{C}_{\text{rad}}$  for which  $\mathcal{G}[\mathbb{T}_{\text{rad}}, M] \leq \mathcal{G}[\mathbb{T}, M]$ .

By Proposition A.1, we can find a set of radial polynomials  $\bar{P}_i$  such that  $\langle \bar{P}_i \rangle = \langle P_i \rangle$  and clearly  $\langle \mathcal{F} \rangle = \langle \mathcal{F}_{\text{rad}} \rangle$ . By symmetry considerations, we know that only radial polynomials with even degree will contribute to the mean value. We now put  $\bar{T}_{i;j} \equiv \bar{T}_{2n} = \frac{\langle P_{2n} \rangle}{m \langle Q^{2n} \rangle}$ . By Proposition A.3, we pick up a set of uniformly distributed cuts,  $\theta_j = \frac{j\pi}{m}$  with  $m$  being, for simplicity, the number of cuts in  $\mathbb{T}$ , so that the identity in Eq. (A2) holds, i.e.,  $\bar{P}_{2n} = \bar{T}_{2n} \sum_j Q_{\theta_j}^{2n}$ . Furthermore, it is seen that both  $T_{i;j}$  and  $\bar{T}_{i;j}$  satisfy the same constraint

$$\sum_{j=1}^m T_{i;j} = \sum_{j=1}^m \bar{T}_{i;j} = \frac{\langle P_{2n} \rangle}{\langle Q^{2n} \rangle}. \quad (\text{A27})$$

Consequently, both  $\mathcal{F} = 1 + \sum_{i,j} T_{i;j} Q_{\theta_j}^i$  and  $\mathcal{F}_{\text{rad}} = 1 + \sum_{i,j} \bar{T}_{i;j} Q_{\theta_j}^i$  satisfy the hypothesis of Proposition A.2 with the same  $K_i = \frac{\langle P_{2n} \rangle}{\langle Q^{2n} \rangle}$ . It follows from the proposition that  $\sigma_{\mathcal{F}_{\text{rad}}}^2$  is the minimum attainable variance within that class or equivalently

$$\sigma_{\mathcal{F}_{\text{rad}}}^2 \leq \sigma_{\mathcal{F}}^2 \quad (\text{A28})$$

Therefore, defining  $\mathbb{T}_{\text{rad}} \equiv (N, \bar{P}_i, \bar{\theta}_j, \bar{T}_{i;j})$  and using that  $\langle \mathcal{F} \rangle = \langle \mathcal{F}_{\text{rad}} \rangle \leq 0$ , we arrive at

$$\mathcal{G}[\mathbb{T}_{\text{rad}}, M] \leq \mathcal{G}[\mathbb{T}, M]. \quad (\text{A29})$$

$\square$

Thus, we have proven that whenever  $W(x, p)$  is rotationally symmetric, a rotationally invariant test is optimal. It should be noted, however, that the test described in the main text is not as general as the one considered in this appendix because the former is always a square polynomial to ensure non-negativity in a simple manner. Moreover, the map we have used

$$P_i \rightarrow \bar{P}_i \quad (\text{A30})$$

does not preserve the property of being a square polynomial for  $\mathcal{F}$ , that is  $\mathcal{F}_{\text{rad}} = 1 + \sum_i \bar{P}_i$  is not generally a square polynomial even when  $\mathcal{F} = 1 + \sum_i P_i$  is. Nonetheless, the square root of  $\mathcal{F}_{\text{rad}}$  can be efficiently approximated by a high order polynomial in  $r^2$  in the relevant region where it is peaked at the negativity of the Wigner function. This, together with the convergence shown in Fig. 2, motivates the choice of polynomials in  $r^2$ .

### 3. Squeezed states.

We will now consider the behaviour of the optimal test for the more general squeezed state.

**Theorem A.2.** *(Invariance under squeezing) Consider the best test  $\mathbb{T}_0 = (N, P_i, \theta_j, T_{i;j}) \in \mathcal{C}$  for the joint distribution  $W(x, p)$ . Then the best test  $\mathbb{T}_\lambda$  for the corresponding squeezed state  $W_\lambda(x, p) = W(\lambda x, p/\lambda)$ ,  $\lambda$  being the squeezing parameter, is characterised by  $\mathbb{T}_\lambda = (N, P_i^{(\lambda)}, \theta_j^{(\lambda)}, T_{i;j}^{(\lambda)})$  where*

$$P_i^{(\lambda)}(x, p) \equiv P_i(\lambda x, p/\lambda), \quad (\text{A31})$$

$$\tan \theta_j^{(\lambda)} \equiv \frac{\tan \theta_j}{\lambda^2}, \quad (\text{A32})$$

$$T_{i;j}^{(\lambda)} \equiv T_{i;j} \left( \lambda \frac{\cos \theta_j}{\cos \theta_j^{(\lambda)}} \right)^i. \quad (\text{A33})$$

*Proof.* The proof is by construction of a bijective correspondence between tests for the unsqueezed case and tests for the squeezed one. First, we consider the representation of the new set of polynomials in terms of the

new set of cuts

$$\begin{aligned} P_i^{(\lambda)}(x, p) &= P_i(\lambda x, p/\lambda) \\ &= \sum_j T_{i;j} \left( \cos \theta_j \lambda x + \sin \theta_j \frac{p}{\lambda} \right)^i \\ &= \sum_j T_{i;j} (\cos \theta_j \lambda)^i \left( x + \frac{\tan \theta_j}{\lambda^2} p \right)^i. \end{aligned} \quad (\text{A34})$$

In the domain  $(-\pi, +\pi)$ , we can find new cuts  $\theta_j^{(\lambda)}$  according to Eq. (A32) and obtain

$$\begin{aligned} P_i^{(\lambda)}(x, p) &= \sum_j T_{i;j} \left( \lambda \frac{\cos \theta_j}{\cos \theta_j^{(\lambda)}} \right)^i \left( \cos \theta_j^{(\lambda)} x + \sin \theta_j^{(\lambda)} p \right)^i \\ &= \sum_j T_{i;j}^{(\lambda)} Q_{\theta_j^{(\lambda)}}^i, \end{aligned} \quad (\text{A35})$$

where  $T_{i;j}^{(\lambda)}$  is defined in Eq. (A33).

We will now show that  $\langle \mathcal{F} \rangle = \langle \mathcal{F}^{(\lambda)} \rangle_{\lambda}$  and that  $\sigma_{\mathcal{F}}^2 = \sigma_{\mathcal{F}^{(\lambda)}}^2$  where  $\mathcal{F}$  ( $\mathcal{F}^{(\lambda)}$ ) is the test function for the unsqueezed (squeezed) state. We recall the result for the Radon transform in Eq. (43) stated below for convenience:

$$\mathbf{RW}_{\lambda}(\theta, s) = \frac{1}{u(\lambda, \theta)} \mathbf{RW} \left( \frac{s}{u(\lambda, \theta)} \right), \quad (\text{A36})$$

where  $\lambda$  is the squeezing parameter and  $u(\lambda, \theta)$  is defined as  $u(\lambda, \theta) = \frac{\cos \theta}{\lambda} \sqrt{1 + \lambda^4 \tan^2 \theta}$ . It follows that

$$\begin{aligned} \langle Q_{\theta}^i \rangle_{\lambda} &\equiv \int_{-\infty}^{+\infty} ds s^i \mathbf{RW}_{\lambda}(\theta, s) \\ &= \int_{-\infty}^{+\infty} \frac{ds}{u(\lambda, \theta)} s^i \mathbf{RW} \left( \frac{s}{u(\lambda, \theta)} \right) \\ &= u^i(\lambda, \theta) \int_{-\infty}^{+\infty} dt t^i \mathbf{RW}(\theta, t) \\ &\equiv u^i(\lambda, \theta) \langle Q^i \rangle, \end{aligned} \quad (\text{A37})$$

where we have changed from the variable  $s$  to  $t \equiv \frac{s}{u(\lambda, \theta)}$ . For the mean values, we find

$$\begin{aligned} \langle P_i^{(\lambda)} \rangle_{\lambda} &\equiv \sum_j T_{i;j}^{(\lambda)} \langle Q_{\theta_j^{(\lambda)}}^i \rangle_{\lambda} \\ &= \sum_j T_{i;j} \left( \lambda \frac{\cos \theta_j}{\cos \theta_j^{(\lambda)}} \right)^i u^i(\lambda, \theta_j^{(\lambda)}) \langle Q^i \rangle. \end{aligned} \quad (\text{A38})$$

Since  $\tan \theta_j^{(\lambda)} \equiv \frac{\tan \theta_j}{\lambda^2}$ , it is seen that  $u^i(\lambda, \theta_j^{(\lambda)}) = \left( \lambda \frac{\cos \theta_j}{\cos \theta_j^{(\lambda)}} \right)^{-i}$  and hence

$$\langle P_i^{(\lambda)} \rangle_{\lambda} = \sum_j T_{i;j} \langle Q^i \rangle \equiv \langle P_i \rangle, \quad (\text{A39})$$

so that  $\langle \mathcal{F} \rangle = \langle \mathcal{F}^{(\lambda)} \rangle_{\lambda}$ .

We now wish to establish a similar result for the variances, i.e.,  $\sigma_{\mathcal{F}^{(\lambda)}}^2 = \sigma_{\mathcal{F}}^2$ . We have that

$$\sigma_{\mathcal{F}^{(\lambda)}}^2 = \frac{1}{2} \sum_j \sum_{i,i'=1}^{2N} g_{\lambda}^{ii'}(\theta_j) T_{i;j}^{(\lambda)} T_{i';j}^{(\lambda)}, \quad (\text{A40})$$

where the covariance matrix  $\frac{1}{2} g_{\lambda}^{ii'}(\theta_j) \equiv \left( \langle Q_{\theta_j}^{i+i'} \rangle_{\lambda} - \langle Q_{\theta_j}^i \rangle_{\lambda} \langle Q_{\theta_j}^{i'} \rangle_{\lambda} \right)$  has been introduced. Eq. (A37) implies that  $g_{\lambda}^{ii'}(\theta_j) = u^{i+i'}(\lambda, \theta_j) g^{ii'}$  so that

$$\sigma_{\mathcal{F}^{(\lambda)}}^2 = \frac{1}{2} \sum_j \sum_{i,i'=1}^{2N} g^{ii'} T_{i;j} T_{i';j} \equiv \sigma_{\mathcal{F}}^2, \quad (\text{A41})$$

because  $u^i(\lambda, \theta_j^{(\lambda)}) = \left( \lambda \frac{\cos \theta_j}{\cos \theta_j^{(\lambda)}} \right)^{-i}$ . Note that the variance for each cut  $\sigma_j^2$  is preserved.

To conclude, we have explicitly constructed a correspondence between any test used in the case of no squeezing  $\mathbb{T}_0$  and a test to use in the squeezed case  $\mathbb{T}_{\lambda}$  such that

$$\mathcal{G}[\mathbb{T}_{\lambda}, M] = \mathcal{G}[\mathbb{T}, M]. \quad (\text{A42})$$

This result establishes that the best test for the squeezed state is at least as good as the one for the unsqueezed state. The transformation applied, however, also works in the opposite direction, so that a more optimal test for the squeezed state, could be unsqueezed and act as a more optimal test for the unsqueezed state. Hence there is a one to one correspondence between the optimal tests between the squeezed and unsqueezed states and the statement of Theorem A.2 follows.  $\square$

## Appendix B: Inverse Radon method

In this appendix, we discuss the results of the filtered back-projection algorithm that have been quoted in the main text without proof. We generalise and prove the result used for the linear transformed kernel (see Eq. (31)), and provide the proof that the distribution of cuts described for the Monte Carlo quadrature method is the optimal distribution. Finally, we discuss the alternative procedure with a finite number of cuts in detail.

The result for the linear transformation of the kernel can be derived for the  $d$ -dimensional case. The planar case ( $d = 2$ ) that has been presented in the main text will then be retrieved as a special case. The whole theory of the Radon transform can be formulated within  $\mathbb{R}^d$  where the Radon transform is achieved by integration over all hyperplanes in  $\mathbb{R}^d$ , rather than straight lines in the plane. For a detailed but not too abstract discussion see for instance Ref. [20]. Here, we only focus on the adjoint Radon transformation. Let  $\mathbf{x}$  denote a vector in  $\mathbb{R}^d$  and  $\langle \cdot | \cdot \rangle$  denote the standard scalar product. A non singular

linear transformation from  $\mathbb{R}^d$  into itself is denoted by  $L$  and we introduce the  $(d-1)$ -dimensional sphere,  $S^{d-1}$ , as the set unit vectors in  $\mathbb{R}^d$ . The measure of the total set  $S^{d-1}$  is

$$|S^{d-1}| \equiv \int_{\mathbf{n} \in S^{d-1}} d\mathbf{n}. \quad (\text{B1})$$

We now give the explicit functional form of the adjoint Radon operator  $\mathbf{R}^\dagger$ , as it can be found in the reference mentioned above. This is an operator that maps kernel-like functions, call them  $\omega(\mathbf{n}, s)$  where  $\mathbf{n} \in S^{d-1}$  and  $s \in \mathbb{R}$ , into (filter-like) functions of  $\mathbf{x} \in \mathbb{R}^d$ , according to

$$\mathbf{R}^\dagger \omega(\mathbf{x}) = \frac{1}{|S^{d-1}|} \int_{\mathbf{n} \in S^{d-1}} \omega(\mathbf{n}, \langle \mathbf{x} | \mathbf{n} \rangle) d\mathbf{n}, \quad (\text{B2})$$

Finally, we define  $\omega_L$ , the linear transformed kernel of  $\omega$  through  $L$ , according to the identity

$$\mathbf{R}^\dagger \omega_L(\mathbf{x}) = \mathbf{R}^\dagger \omega(L(\mathbf{x})). \quad (\text{B3})$$

If  $\omega$  is the kernel for the filter  $F$ , *i.e.*,  $\mathbf{R}^\dagger \omega(\mathbf{x}) = F(\mathbf{x})$ , then  $\omega_L$  is the kernel for  $F(L\mathbf{x})$ . This is precisely the situation encountered for the squeezed filter provided that  $L$  is the squeezing transformation. The following exact result holds.

**Theorem B.1.** *(Behaviour under linear transformation)* *Given the filter  $\omega(\mathbf{n}, s)$  and an invertible linear transformation  $L$ , then the linear transformed kernel  $\omega_L(\mathbf{n}', s)$  is given by*

$$\omega_L(\mathbf{n}', s) = |\det \mathbb{J}_\phi(\mathbf{n})|^{-1} \omega(\mathbf{n}, |L^T \mathbf{n}| s). \quad (\text{B4})$$

We have here put  $\mathbf{n}' = \phi(\mathbf{n}) \equiv \frac{L^T \mathbf{n}}{|L^T \mathbf{n}|}$ , which is invertible on the semisphere, and  $\mathbb{J}_\phi$  is its Jacobian matrix. Note that the substitution  $\mathbf{n} = \phi^{-1}(\mathbf{n}')$  is intended in the formula above.

*Proof.* The application of the adjoint operator to the identity that defines  $\omega_L(x)$  (B3) leads to

$$\mathbf{R}^\dagger \omega(L\mathbf{x}) = \frac{1}{|S^{d-1}|} \int_{\mathbf{n} \in S^{d-1}} \omega(\mathbf{n}, \langle L\mathbf{x} | \mathbf{n} \rangle) d\mathbf{n}. \quad (\text{B5})$$

Considering the identity  $\langle L\mathbf{x} | \mathbf{n} \rangle = \langle \mathbf{x} | L^T \mathbf{n} \rangle$  with  $L^T$  denoting the transposed transformation, it is natural to change the variable of the integration to  $\mathbf{n}' = \phi(\mathbf{n}) \equiv \frac{L^T \mathbf{n}}{|L^T \mathbf{n}|}$ ,  $\mathbf{n}' \in S^{d-1}$ , which leads to

$$\begin{aligned} \mathbf{R}^\dagger \omega_L(\mathbf{x}) &= \mathbf{R}^\dagger \omega(L\mathbf{x}) \\ &= \frac{1}{|S^{d-1}|} \int_{\mathbf{n}' \in S^{d-1}} \omega(\mathbf{n}, |L^T \mathbf{n}| \langle \mathbf{x} | \mathbf{n}' \rangle) \frac{d\mathbf{n}'}{|\det \mathbb{J}_\phi(\mathbf{n})|}, \end{aligned} \quad (\text{B6})$$

which is the adjoint Radon transform of the function  $\omega_L(\mathbf{n}', s)$ . The statement of Theorem B.1 directly follows by putting  $\langle \mathbf{x} | \mathbf{n}' \rangle = s$ .  $\square$

For our application of the theorem, we consider a phase plane,  $\mathbb{R}^2$ , and the linear transformation  $L$  corresponds to a squeezing transformation (the case of dilation, *i.e.*  $L$  is a multiple of the identity is trivial). We write  $\mathbf{n}(\theta) = (\cos \theta, \sin \theta)$  and recover the notation used for the kernel in the main text, *i.e.*,  $\omega(\theta, s) \equiv \omega(\mathbf{n}(\theta), s)$ . The following corollary holds.

**Corollary B.1.** *The general filter that can be obtained from a rotational invariant one, denoted by  $\omega(s)$ , by dilation and rescaling is of the form*

$$\omega_{a,\lambda}(\theta, s) = \frac{1}{u^2(\lambda, \theta)} \omega_a \left( \frac{s}{u(\lambda, \theta)} \right), \quad (\text{B7})$$

with  $\omega_a(s)$  defined in Eq. (30) so that  $\omega_{a,\lambda}(\theta, s)$  is normalized as well.  $\lambda$  is the rescaling parameter.

*Proof.* We parametrize the rescaling transformation by  $L(\lambda) = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$ . We also introduce the the following parametrizations  $\mathbf{n}(\theta) = (\cos \theta, \sin \theta)$  and  $\mathbf{n}'(\theta') = (\cos \theta', \sin \theta')$ . Then one can write

$$\begin{aligned} \tan \theta' &= \lambda^{-2} \tan \theta \\ |L^T \mathbf{n}| &= u^{-1}(\theta', \lambda) \\ |\det \mathbb{J}_\phi(\mathbf{n})|^{-1} &= \frac{1 + \tan^2 \theta'}{\lambda^{-2} + \tan^2 \theta'} = |L^T \mathbf{n}|^2 \end{aligned} \quad (\text{B8})$$

and by substitution the proof is concluded. Note the simplification in notation  $\omega_{a,\lambda}(\theta, s) \equiv \omega_{a,L(\lambda)}(\theta, s)$ .  $\square$

At this point, we have established the general results on the filter back-projection that we have used throughout the application of the Inverse Radon method. We now turn to the particular problem of the determining the best distribution of cuts  $\mathcal{C}(\theta)$ , (see Eq. (37)), for the Monte Carlo quadrature procedure.

## 1. Monte Carlo quadrature.

By definition, the mean value of the test function and  $\Delta_a(\epsilon)$  do not depend on  $\mathcal{C}(\theta)$  nor on  $\lambda$  [see Eqs. (44) and (35)]. Nonetheless, the variance of the estimate  $\langle F_{a,\lambda} \rangle$  has a strong dependence on  $\mathcal{C}(\theta)$ , that is

$$\begin{aligned} \sigma_{F_{a,\lambda}}^2[\mathcal{C}; \epsilon, M] &\equiv \\ &\equiv \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \int_{\mathbb{R}} p_\lambda(\theta, s) \frac{(I_{a,\lambda,\epsilon}(\theta, s) - \langle F_{a,\lambda} \rangle)^2}{M-1} ds d\theta \\ &\equiv \frac{\mathcal{W}_{a,\epsilon}}{M-1} \left( \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{d\theta}{\mathcal{C}(\theta) u^4(\lambda, \theta)} \right) - \frac{\langle F_{a,\lambda} \rangle^2}{M-1}, \end{aligned} \quad (\text{B9})$$

where the quantity  $\mathcal{W}_{a,\epsilon} \equiv \left( \int_{\mathbb{R}} \mathbf{R}W(t) \frac{\omega_{a,\epsilon}^2(t)}{\pi^2} dt \right)$  has been defined. Note that this quantity only relates to the unsqueezed distribution and all the  $\lambda$ -dependence of the variance  $\sigma_{F_{a,\lambda}}^2$  is contained in the integral  $\left( \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{d\theta}{\mathcal{C}(\theta) u^4(\lambda, \theta)} \right)$ , which also depends upon the distribution  $\mathcal{C}$ . Since our present concern is a suitable choice

of  $\mathcal{C}$ , the integral in the expression above is the main focus of the analysis. In other words, for fixed  $a$ ,  $\lambda$ ,  $\epsilon$  and  $M$ , we wish to determine the distribution  $\mathcal{C}$  that minimises  $\left(\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{d\theta}{\mathcal{C}(\theta)u^4(\lambda,\theta)}\right)$  and  $\sigma_{F_{a,\lambda}}^2[\mathcal{C};\epsilon,M]$  accordingly. We can now state the following exact result concerning the relevant ratio in Eq. (40).

**Corollary B.2.** *Assume the state to be described by a squeezed distribution  $W_\lambda(x,p)$ , where for  $\lambda = 1$  the distribution is rotationally invariant. Then the optimal distribution of cuts  $\mathcal{C}_o$ , which minimises the relevant ratio (40), only depends on  $\lambda$  through*

$$\mathcal{C}_o(\theta) = \frac{1}{\pi u^2(\lambda,\theta)}. \quad (\text{B10})$$

Moreover, for this choice of phase cut distribution the value of the relevant ratio does not depend on  $\lambda$  and it is thus equal to the one for the rotationally invariant state ( $\lambda = 1$ ).

*Proof.* We have already pointed out that the dependence on the distribution of cuts in the ratio  $\frac{\langle F_{a,\lambda} \rangle}{\sigma_{F_{a,\lambda}}^2 + \Delta_a(\epsilon)}$  is only through the integral  $\left(\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{d\theta}{\mathcal{C}(\theta)u^4(\lambda,\theta)}\right)$  appearing in the variance. We are thus considering a constrained optimisation problem and can define the Lagrangian functional

$$\mathcal{L}[\mathcal{C},\mu] \equiv \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \left( \frac{1}{\mathcal{C}(\theta)u^4(\lambda,\theta)} + \mu \mathcal{C}(\theta) \right) d\theta. \quad (\text{B11})$$

together with the normalisation condition,  $\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \mathcal{C}(\theta) d\theta = 1$ , that determines the Lagrange multiplier  $\mu \in \mathbb{R}$ . Let us consider a small variation of the distribution of cuts,  $\mathcal{C}(\theta) \rightarrow \mathcal{C}(\theta) + \delta\mathcal{C}(\theta)$ . Then, the stationary point equation is given by

$$0 = \delta\mathcal{L}[\mathcal{C},\mu] = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \left( -\frac{1}{\mathcal{C}^2(\theta)u^4(\lambda,\theta)} + \mu \right) \delta\mathcal{C}(\theta) d\theta, \quad (\text{B12})$$

where  $\delta\mathcal{L}[\mathcal{C},\mu]$  is the linear part of  $\mathcal{L}[\mathcal{C} + \delta\mathcal{C},\mu] - \mathcal{L}[\mathcal{C},\mu]$  with respect to  $\delta\mathcal{C}$ . The solution of Eq. (B12) is the stationary distribution of cuts  $\mathcal{C}_o = \frac{1}{\sqrt{\mu} u^2(\lambda,\theta)}$ . By integration,

$$\begin{aligned} \int_{-\pi/2}^{\pi/2} \frac{d\theta}{u^2(\lambda,\theta)} &= \int_{-\pi/2}^{\pi/2} \frac{\lambda^2}{\cos^2 \alpha} \frac{d\alpha}{1 + \lambda^4 \tan^2 \alpha} \\ &= \int_{-\infty}^{+\infty} \frac{dx}{1 + x^2} = \pi, \end{aligned} \quad (\text{B13})$$

we find the normalization condition  $\mu = \pi^2$ . We can now check that the stationary distribution  $\mathcal{C}_o$  gives the minimum of the Lagrangian, by considering the second variation around  $\mathcal{C}_o$ , that is

$$\delta^2 \mathcal{L}[\mathcal{C}_o,\mu] = \pi^3 \mathcal{W}_a \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \delta\mathcal{C}^2(\theta) u^2(\lambda,\theta) d\theta > 0, \quad (\text{B14})$$

which is positive for any variation  $\delta\mathcal{C}$ .  $\mathcal{C}_o$  thus gives the minimum. Finally, inserting  $\mathcal{C}_o$  into  $\sigma_{F_{a,\lambda}}^2[\mathcal{C};\epsilon,M]$  shows that the resulting variance  $\sigma_{F_{a,\lambda}}^2[\mathcal{C};\epsilon,M]$  is independent of  $\lambda$ . This completes the proof of Corollary B.2.  $\square$

The naive choice of uniformly distributed cuts  $\mathcal{C} = \frac{1}{\pi}$  leads to a poor variance for a general squeezed state. This is easily seen by straightforward computation of the integral

$$\int_{-\pi/2}^{+\pi/2} \frac{d\theta}{u^4(\theta,\lambda)} = \int_{-\pi/2}^{+\pi/2} \frac{\lambda^4}{\cos^4 \alpha} \frac{d\alpha}{(1 + \lambda^4 \tan^2 \alpha)^2} \quad (\text{B15})$$

$$= \int_{-\infty}^{+\infty} \frac{\lambda^2 + \lambda^{-2} x^2}{(1 + x^2)^2} dx = \frac{\pi}{2} (\lambda^2 + \lambda^{-2}), \quad (\text{B16})$$

so that

$$\sigma_{F_{a,\lambda}}^2 = \frac{\mathcal{W}_{a,\epsilon}}{M-1} \frac{(\lambda^2 + \lambda^{-2})}{2} - \frac{\langle F_{a,\lambda} \rangle^2}{M-1}. \quad (\text{B17})$$

The first term is dominant for  $\lambda \gg 1$  ( $\lambda \ll 1$ ), which gives the scaling of  $\sim \lambda^2$  ( $\sim \lambda^{-2}$ ).

## 2. Finite number of cuts.

Here, we consider the details of the inverse Radon method with a finite number of cuts in phase space. As remarked in the main text, we assume that the states are described by a symmetric  $W_\lambda(x,p)$  phase space distribution, where for  $\lambda = 1$  this is the rotationally symmetric single photon state. We wish to classify the best way to carry out the approximation

$$\langle F_{a,\lambda} \rangle \simeq \sum_{\{\theta_j\}} w_j g_{a,\lambda}(\theta_j) \quad (\text{B18})$$

with only a limited number of phase cuts. By the symmetry assumed, the mean value of the test function is (see Eqs. (31,43,44))

$$\begin{aligned} \langle F_{a,\lambda} \rangle &= \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{d\theta}{u^2(\lambda,\theta)} \left( \int_{\mathbb{R}} \mathbf{R} W_{\lambda=1}(t) \omega_{a,\lambda=1}(t) dt \right) \\ &= \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} d\theta \frac{1}{u^2(\lambda,\theta)} g_{a,\lambda=1}. \end{aligned} \quad (\text{B19})$$

Here, it is important that  $g_{a,\lambda=1}$  does not depend on the phase. The problem thus concerns how to carry out the integration of  $\frac{1}{u^2(\theta,\lambda)}$  numerically. In other words, if  $g_{a,\lambda=1}$  does not contain any experimental imprecision, the only error associated to the estimate  $\langle F_{a,\lambda} \rangle$  comes from the discretisation of  $\int \frac{d\theta}{u^2(\theta,\lambda)}$ . This problem can be addressed with use of the composite middle-point formula. The numerical quadrature is obtained by choosing a partition of the interval  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  into non-uniform

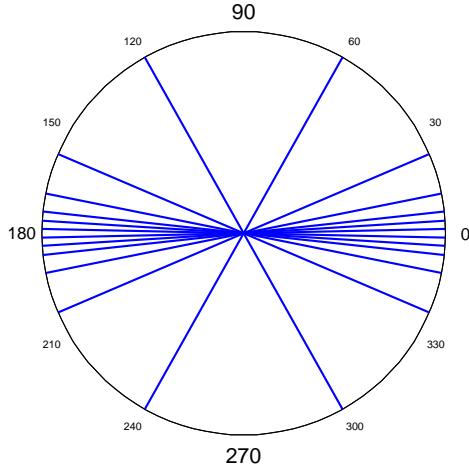


FIG. 8 (colour online). Optimal position for 12 cuts in the middle-point quadrature. Here  $\lambda = 5$  and the precision of the numerical quadrature is within the 3% of error.

subintervals, in which the value of the integral is approximated by the product of the length of the subinterval times the value of the integrand in its middle point. We exploit the parity of the integrand and choose partition points symmetric around  $\theta = 0$ , the latter included. The complete integral is then approximated as

$$\frac{1}{\pi} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} d\theta \frac{1}{u^2(\lambda, \theta)} \simeq \frac{1}{\pi} \sum_{j=1}^m \frac{\Delta\theta_j}{u^2(\tilde{\theta}_j, \lambda)}, \quad (\text{B20})$$

where  $\tilde{\theta}_j$  and  $\Delta\theta_j$  are the middle point and the size of the  $j$ -th subinterval, respectively. The middle points should correspond to the cuts where measurements are carried out. In analogy with the elementary test,  $m$  is the total number of cuts over the entire interval  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ . By comparison with Eq. (B18), the weights are  $w_j = \frac{\Delta\theta_j}{\pi}$ . The evaluation at the middle point is a convenient choice because the integrand is monotonically decreasing (increasing) within the interval  $[0, +\frac{\pi}{2}]$  ( $[-\frac{\pi}{2}, 0]$ ). The middle-point evaluation thus provides a trade off between the surplus in one half of the interval and the deficiency in the other.

We define the numerical error to be

$$\mathcal{E} \equiv \left| 1 - \frac{1}{\pi} \sum_{j=1}^m \frac{\Delta\theta_j}{u^2(\tilde{\theta}_j, \lambda)} \right|. \quad (\text{B21})$$

For fixed  $\lambda$  and fixed  $m$ , we have numerically optimised  $\mathcal{E}$  with respect to the position of the cuts. As an example, we show in Fig. 8 the result for  $\lambda = 5$  and  $m = 12$ . We observe a concentration of cuts near  $\theta = 0$ , in analogy with the elementary test. Moreover, as shown in Fig. 9,  $\mathcal{E}$  decreases roughly exponentially as a function of  $m$ . The error can, in principle be made arbitrarily small by increasing the number of cuts. The bigger  $\lambda$  is, however, the more cuts are needed to push the error

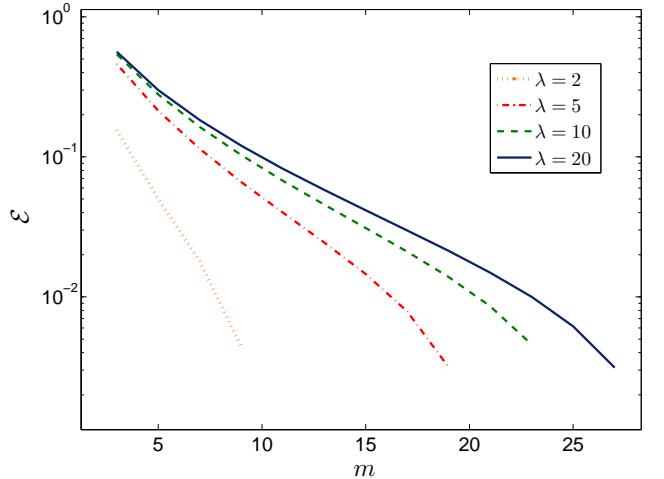


FIG. 9 (colour online). The numerical error  $\mathcal{E}$  as a function of the number of cuts  $m$  for various values of the squeezing parameter  $\lambda$ . The plot has been truncated at the onset of numerical instabilities that causes the error to oscillate around  $10^{-5}$ . Notice the nearly exponential decrease for each curve and that for larger  $\lambda$  more cuts are needed to reach a given error.

below a given threshold. Hence, in an actual experiment with a squeezed state, it will be necessary to use a large number of cuts to constrain the error from the numerical quadrature. Also, one would have to justify using a highly asymmetric distribution of cuts to estimate an integral over all angles. This can be argued from the asymmetry of the filter, which is the main reason for strong angular dependence, but such an argument would still leave open the question of whether the state itself have rapid angular variations in the region with a less dense distribution of cuts. Alternatively, one would have to resort to very small intervals between the cuts for all angles. If one is using the same number of measurements per cut, this would be very costly in terms of measurements since a large number of measurements are distributed in a region with little dependence on the angle. One can to some degree compensate for this with an asymmetric distribution of the number of measurement (see below), but still the issue of the precision of the numerical quadrature is an uncomfortable question, which it would be desirable to avoid, *e.g.*, by using the elementary test.

Note also that, for simplicity, the discussion above has been carried out by assuming exact knowledge of  $g_{a,\lambda}$ . In an actual experiment, where we want to present a proof of non-classicality with as few assumptions as possible, we do not want to rely on a specific theoretical model for the underlying distribution. Instead, one would have to justify a specific bound on the error coming from having a finite number of cuts by considering the actual measured experimental data. This is significantly more complicated than the analysis performed here, and we shall not discuss it.

In addition to the error from the finite number of cuts,

we also we have to consider the statistical error. If the cuts are considered to be statistically independent, the variance of the estimate  $\langle F_{a,\lambda} \rangle$  is

$$\sigma_{F_{a,\lambda}}^2 = \sum_{\{\theta_j\}} \frac{W_j^2}{M_j - 1} \sigma_a^2, \quad (\text{B22})$$

where for brevity we have put  $W_j = \frac{w_j}{u^2(\theta_j, \lambda)}$ , satisfying  $\sum_j W_j = 1$ . Note that the  $\lambda$  dependence is implicit in the optimal  $\{\theta_j\}$  and  $W_j$ . Also, the variance of  $g_a(\theta_j)$  is independent of  $\theta_j$ . For uniform weights, it is desirable to have the same number of measurements  $M_j$  for each cut. For the optimal quadrature, where  $W_j$  are not uniform, this is no longer true, since a small weight  $W_j$  allows to tolerate a rather big imprecision in  $g_a(\theta_j)$ . Given the total number of measurement  $M$  and  $m$  cuts, the optimal distribution of measurements per cut  $M_j$  can be obtained analytically. Formally, the problem consists in the optimisation of the quantity in Eq. (B22) with respect to  $M_j$ , constrained by  $\sum_j (M_j - 1) = M - m$ . The Lagrangian for this optimisation problem is

$$\mathcal{L}(M_j, \mu) = \sigma_a^2 \sum_{j=1}^m \frac{W_j^2}{M_j - 1} + \mu \sum_{j=1}^m (M_j - 1), \quad (\text{B23})$$

which has the optimal solution

$$\begin{cases} M_j - 1 = \frac{\sigma_a W_j}{\sqrt{\mu}} \\ \mu = \left( \frac{\sigma_a \sum_j W_j}{M - m} \right)^2 \end{cases} \quad (\text{B24})$$

By eliminating the Lagrange multiplier  $\mu$ , we find  $(M_j - 1) = \frac{(M - m) W_j}{\sum_j W_j}$  and the minimal value of the variance is

$$\sigma_{F_{a,\lambda}}^2 = \frac{\left( \sum_{j=1}^m W_j \right)^2}{M - m} \sigma_a^2 = \frac{\sigma_a^2}{M - m}. \quad (\text{B25})$$

The last equality follows from  $\sum_{\{\theta_j\}} W_j = 1$  and establishes the independence on the squeezing, similar to the results obtained for the other methods. Notice that we have considered  $M_j \in \mathbb{R}$  for simplicity; the fact that they are natural numbers introduce a small deviation from the optimum determined here.

The total error associated with the estimate  $\langle F_{a,\lambda} \rangle$  is given by the sum of all contributions

$$\sigma_{F_{a,\lambda}}^2 + \mathcal{E} + \Delta_a(\epsilon). \quad (\text{B26})$$

Similar to the other test this can be made independent of  $\lambda$  provided that  $m$  is big enough (see also Fig. 9 and discussion; notice although that a large  $m$  can decrease the strength of the test if  $M$  is not very big). Due to the large number of ambiguities which would have to be resorted in order to make a sensible comparison with the elementary test, we will, however, not make a more detailed comparison.

---

[1] P. Dirac, *Principles of Quantum Mechanics* (SIAM Society for Industrial and Applied Mathematics, 1926)

[2] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (1932)

[3] 't Hooft Gerard, arXiv preprint arXiv:14051548v2.0885(Jun 2014), [14051548v2.0885 \[quant-ph\]](https://arxiv.org/abs/1405.1548v2.0885)

[4] J. Fröhlich and B. Schubnel, arXiv preprint arXiv:1203.3678(2012)

[5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information (Cambridge Series on Information and the Natural Sciences)* (Cambridge University Press, Cambridge, 2004, 2004)

[6] E. Kot, N. Grønbech-Jensen, B. M. Nielsen, J. S. Neergaard-Nielsen, E. S. Polzik, and A. S. Sørensen, *Phys. Rev. Lett.* **108**, 233601 (Jun 2012), [http://link.aps.org/doi/10.1103/PhysRevLett.108.233601](https://link.aps.org/doi/10.1103/PhysRevLett.108.233601)

[7] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, 1st ed. (Pisa: Edizioni della Normale, 2011)

[8] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, *Rev. Mod. Phys.* **86**, 419 (Apr 2014), [http://link.aps.org/doi/10.1103/RevModPhys.86.419](https://link.aps.org/doi/10.1103/RevModPhys.86.419)

[9] F. Marquardt and S. M. Girvin, *Physics* **2**, 40 (May 2009), [http://link.aps.org/doi/10.1103/Physics.2.40](https://link.aps.org/doi/10.1103/Physics.2.40)

[10] M. Aspelmeyer, S. Gröblacher, K. Hammerer, and N. Kiesel, *J. Opt. Soc. Am. B* **27**, A189 (Jun 2010), <http://josab.osa.org/abstract.cfm?URI=josab-27-6-A189>

[11] K. Hammerer, A. S. Sørensen, and E. S. Polzik, *Rev. Mod. Phys.* **82**, 1041 (Apr 2010), [http://link.aps.org/doi/10.1103/RevModPhys.82.1041](https://link.aps.org/doi/10.1103/RevModPhys.82.1041)

[12] J. Wenger, M. Hafezi, F. Grosshans, R. Tualle-Brouri, and P. Grangier, *Phys. Rev. A* **67**, 012105 (Jan 2003), [http://link.aps.org/doi/10.1103/PhysRevA.67.012105](https://link.aps.org/doi/10.1103/PhysRevA.67.012105)

[13] J. Glauber, *Quantum Theory of Optical Coherence* (WILEY-VCH Verlag GmbH and Co. KGaA, 2007)

[14] W. V. D.-G. Welsch and T. Opatrny, *Progress in Optics* **39**, 63 (1999)

[15] W. Vogel, *Phys. Rev. Lett.* **84**, 1849 (Feb 2000), [http://link.aps.org/doi/10.1103/PhysRevLett.84.1849](https://link.aps.org/doi/10.1103/PhysRevLett.84.1849)

[16] A. I. Lvovsky and J. H. Shapiro, *Phys. Rev. A* **65**, 033830 (Feb 2002), [http://link.aps.org/doi/10.1103/PhysRevA.65.033830](https://link.aps.org/doi/10.1103/PhysRevA.65.033830)

[17] E. Shchukin, T. Richter, and W. Vogel, *Journal of Optics B: Quantum and Semiclassical Optics* **6**, S597 (2004), [http://stacks.iop.org/1464-4266/6/i=6/a=020](https://stacks.iop.org/1464-4266/6/i=6/a=020)

[18] A. Bednorz and W. Belzig, *Phys. Rev. A* **83**,

052113 (May 2011), <http://link.aps.org/doi/10.1103/PhysRevA.83.052113>

[19] J. Korbicz, J. Cirac, J. Wehr, and M. Lewenstein, *Phys. Rev. Lett.* **94**, 153601 (Apr 2005), <http://link.aps.org/doi/10.1103/PhysRevLett.94.153601>

[20] S. Deans, *The Radon Transform and some Its Applications* (A Wiley-Interscience Publication, 1983)

[21] A. Markoe, *Analytic Tomography, Encyclopedia of Mathematics and its Applications* (Cambridge Uni. Press, 2006)

[22] Y. Nievergelt, *Journal of Math. Anal. and Appl.* **120**, 288 (1986)

[23] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, *Rev. Mod. Phys.* **86**, 1391 (Dec 2014), <http://link.aps.org/doi/10.1103/RevModPhys.86.1391>